Mathematical analysis of models of congested road traffic

Roméo Hatchi

Université Paris-Dauphine

N° attribué par la bibliothèque

THÈSE

pour obtenir le grade de

Docteur en sciences de l'Université Paris-Dauphine

Spécialité : **Mathématiques**

préparée au laboratoire **CEREMADE**

dans le cadre de l' **École Doctorale de Dauphine**

présentée et soutenue publiquement
par

Roméo Hatchi

le 02 décembre 2015

Titre:
Analyse mathématique de modèles de trafic routier congestionné

Directeur de thèse: **Guillaume Carlier**

Jury

M. Yves Achdou,	Rapporteur
M. Jean-David Benamou,	Examinateur
M. Pierre Cardaliaguet,	Examinateur
M. Bruno Nazaret,	Examinateur
M. Filippo Santambrogio,	Examinateur
M. Guillaume Carlier,	Directeur de thèse

Après avis de Yves Achdou (rapporteur) et Roberto Cominetti (rapporteur).

Résumé

Cette thèse est dédiée à l'étude mathématique de quelques modèles de trafic routier congestionné. La notion essentielle est l'équilibre de Wardrop. Elle poursuit des travaux de Carlier et Santambrogio avec des coauteurs. Baillon et Carlier [12] ont étudié le cas de grilles cartésiennes dans $\mathbb{R}^2$ de plus en plus denses, dans le cadre de la théorie de Γ-convergence. Trouver l'équilibre de Wardrop revient à résoudre des problèmes de minimisation convexe.

Dans le chapitre 2, nous regardons ce qui se passe dans le cas de réseaux généraux, de plus en plus denses, dans $\mathbb{R}^d$. Des difficultés nouvelles surgissent par rapport au cas initial de réseaux cartésiens et pour les contourner, nous introduisons la notion de courbes généralisées. Des hypothèses structurelles sur ces suites de réseaux discrets sont nécessaires pour s'assurer de la convergence. Cela fait alors apparaître des fonctions qui sont des sortes de distances de Finsler et qui rendent compte de l'anisotropie du réseau. Nous obtenons ainsi des résultats similaires à ceux du cas cartésien.

Dans le chapitre 3, nous étudions le modèle continu et en particulier, les problèmes limites. Nous trouvons alors des conditions d'optimalité à travers une formulation duale qui peut être interprétée en termes d'équilibres continus de Wardrop. Cependant, nous travaillons avec des courbes généralisées et nous ne pouvons pas appliquer directement le théorème de Prokhorov, comme cela a été le cas dans [12, 39]. Pour pouvoir néanmoins l'utiliser, nous considérons une version relaxée du problème limite, avec des mesures d'Young.

Dans le chapitre 4, nous nous concentrons sur le cas de long terme, c'est-à-dire, nous fixons uniquement les distributions d'offre et de demande. Comme montré dans [31], le problème de l'équilibre de Wardrop est équivalent à un problème à la Beckmann et il se réduit à résoudre une EDP elliptique, anisotropique et dégénérée. Nous utilisons la méthode de résolution numérique de Lagrangien augmenté présentée dans [21] pour proposer des exemples de simulation.

Enfin, le chapitre 5 a pour objet l'étude de problèmes de Monge avec comme coût une distance de Finsler. Cela se reformule en des problèmes de flux minimal et une discrétisation de ces problèmes mène à un problème de point-selle. Nous le résolvons alors numériquement, encore grâce à un algorithme de Lagrangien augmenté.

Mots-clés : problème de Monge, trafic congestionné, équilibre de Wardrop, Γ-convergence, courbes généralisées, conditions d'optimalité, mesure d'Young, problème de Beckmann, EDPs anisotropiques et dégénérées, Lagrangien augmenté, simulations numériques, distance de Finsler.

Abstract

This thesis is devoted to the mathematical analysis of some models of congested road traffic. The essential notion is the Wardrop equilibrium. It continues Carlier and Santambrogio's works with coauthors. With Baillon [12] they studied the case of two-dimensional cartesian networks that become very dense in the framework of Γ-convergence theory. Finding Wardrop equilibria is equivalent to solve convex minimisation problems.

In Chapter 2 we look at what happens in the case of general networks, increasingly dense. New difficulties appear with respect to the original case of cartesian networks. To deal with these difficulties we introduce the concept of generalized curves. Structural assumptions on these sequences of discrete networks are necessary to obtain convergence. Sorts of Finsler distance are used and keep track of anisotropy of the network. We then have similar results to those in the cartesian case.

In Chapter 3 we study the continuous model and in particular the limit problems. Then we find optimality conditions through a duale formulation that can be interpreted in terms of continuous Wardrop equilibria. However we work with generalized curves and we cannot directly apply Prokhorov's theorem, as in [12, 39]. To use it we consider a relaxed version of the limit problem with Young's measures.

In Chapter 4 we focus on the long-term case, that is, we fix only the distributions of supply and demand. As shown in [31] the problem of Wardrop equilibria can be reformulated in a problem à la Beckmann and reduced to solve an elliptic anisotropic and degenerated PDE. We use the augmented Lagrangian scheme presented in [21] to show a few numerical simulation examples.

Finally Chapter 5 is devoted to studying Monge problems with as cost a Finsler distance. It leads to minimal flow problems. Discretization of these problems is equivalent to a saddle-point problem. We then solve it numerically again by an augmented Lagrangian algorithm.

Keywords : Monge problem, congested traffic, Wardrop equilibrium, Γ-convergence, generalized curves, optimality conditions, Young's measure, Beckmann problem, anisotropic and degenerated PDEs, augmented Lagrangian, numerical simulations, Finsler distance.

Remerciements

Je souhaite remercier très chaleureusement Guillaume Carlier, qui a dirigé cette thèse durant ces trois années, pour sa patience et sa disponibilité, en particulier au Canada. Je lui suis très reconnaissant de m'avoir transmis son intérêt pour le transport optimal et de m'avoir fait découvrir de nombreux problèmes mathématiques avec pédagogie. Ses conseils et ses commentaires m'ont aidé de nombreuses fois.

Je suis également reconnaissant envers Yves Achdou et Roberto Cominetti qui ont accepté d'être rapporteurs de ma thèse pour l'intérêt qu'ils ont porté à ma thèse. Un grand merci à Jean-David Benamou, Pierre Cardaliaguet, Bruno Nazaret et Filippo Santambrogio d'avoir accepté de faire partie du jury.

Merci particulièrement à Jean-David Benamou pour des discussions fructueuses. Son aide fut précieuse.

Des remerciements vont tout naturellement à l'ensemble des membres du CEREMADE, passés et présents, où l'ambiance décontractée qui règne est propice au travail mathématique. Je remercie les membres du secrétariat Marie Belle, Isabelle Bellier et César Faivre pour leur aide. Un grand merci à Valérie Derouiche pour sa disponibilité et pour m'avoir toujours mis dans de bonnes conditions de travail.

Je n'oublie pas de mentionner les deux interprètes qui m'ont suivi pendant ces 3 années. Merci à Anne-Edmée et Cindy. Avec le temps, elles sont devenues des amies.

Je tiens à remercier mes collègues. Un merci particulier à Rémi pour les nombreuses discussions et les parties de squash et Maxime pour ses relectures et les voyages faits ensemble. Merci à tous les doctorants du CEREMADE avec qui j'ai passé de bons moments. Je pense notamment à Sofia, Viviana, Isabelle, Roxana, Clara, Jessica, Fang, Raphaël, Thibaut, Lenaïc, Camille et Marco. J'en profite également pour remercier les membres du département de mathématiques et statistiques de l'université de Victoria, qui ont rendu mon séjour canadien très agréable, surtout Martial Agueh.

Je souhaite également remercier Yuxin Ge, mon encadrant de stage de Master 2, et qui m'a suggéré de faire mon doctorat avec Guillaume Carlier.

Je voudrais également saluer tous mes amis, les parisiens, les lyonnais, les marseillais, les matheux, les non-matheux, les copains de vacances et, plus généralement, les gens que je vois régulièrement depuis des années : Bruno, Ahmed, Rémy, Simon, Yovann, Alexandre, Pif, Serge, Léa, Mathilde, Delphine, Marion, Matthieu, Emmanuelle, Alexandre, Clément, Adrien, Mehdi, Oliver, Lucie, Laure, Alex, Théo,... auxquels il faudra bien sûr rajouter tous ceux que j'ai oubliés.

Enfin, un dernier mot pour la famille. Merci à mes parents et ma soeur pour leurs encouragements et leur aide. Ils ont toujours été là et sont pour beaucoup dans ce travail. Merci aussi à mes grand-parents.

Table des matières

Table des figures

Chapitre 1

Introduction générale

1 Résultats antérieurs et présentation du problème

1.1 Problèmes de Monge-Kantorovich

Le premier problème de transport optimal a été présenté par Monge en 1781 dans son célèbre mémoire [81]. Le problème consistait à trouver le meilleur moyen de déplacer un tas de sable (*déblais*) vers un trou (*remblais*) en effectuant le moins de travail possible. Plus précisément, en langage mathématique moderne, étant données deux mesures de probabilité μ et ν définies sur $\mathbb{R}^d$ (nous supposons que le tas et le trou ont la même masse, qui est égale à 1). En notant $\mathrm{supp}(\mu) = X$ et $\mathrm{supp}(\nu) = Y$, nous cherchons une application $T : X \mapsto Y$ telle que $\nu = T_{\#}\mu$, c'est-à-dire, T envoie μ sur ν :

$$\int_X h(T(x))\, d\mu(x) = \int_Y h(y)\, d\nu(y)$$

pour toute fonction h continue et bornée sur $\mathbb{R}^d$, et qui minimise la quantité

$$I(T) := \int_{\mathbb{R}^d} |T(x) - x|\, d\mu(x)$$

parmi toutes les applications admissibles.

Plus généralement, considérons la fonction de coût $c : X \times Y \mapsto \mathbb{R}_+$ telle que $c(x, y)$ représente le travail nécessaire pour déplacer une unité de masse de la position $x \in X$ à une nouvelle position $y \in Y$. Le problème de Monge devient :

$$\inf \left\{ I(T) := \int_X c(x, T(x))\, d\mu(x) : \nu = T_{\#}\mu \right\} \tag{1.1}$$

Dans le problème initial de Monge, le travail était proportionnel à la distance parcourue : $c(x, y) = |x - y|$ tout simplement. C'était un problème mathématique difficile, notamment à cause de la forme hautement non linéaire de la contrainte. Ainsi, si μ et ν admettent respectivement des fonctions de densité f et g assez régulières et si T est injective, la condition $\nu = T_{\#}\mu$ devient, après un changement de variables, une EDP de la forme suivante :

$$f(x) = g(T(x))|\det(DT(x))|.$$

Une question naturelle pour le problème (1.1) est de chercher comment prouver l'existence d'un élément minimisant. Une idée évidente est de prendre une suite minimisante et de trouver une bonne limite. Cependant, on ne peut pas utiliser des arguments classiques de compacité (quelle que soit la topologie utilisée) car la non-linéarité de la condition empêche de montrer que la limite est une solution admissible. Le problème a été traité en partie par Appell dans son mémoire [11] mais il a été loin de répondre à toutes les questions, notamment sur l'existence d'une solution optimale et sa caractérisation.

Dans les années 1940, Kantorovich [70, 71] a eu l'idée d'introduire une variante relaxée du problème (1.1) et d'utiliser un principe de dualité pour la programmation linéaire. Tout d'abord, il considéra les mesures de probabilité π sur $X \times Y$ telles que

$$\pi(A \times Y) = \mu(A) \ \text{ et } \ \pi(X \times B) = \nu(B) \tag{1.2}$$

pour tous sous-ensembles boréliens A de X et B de Y. Notons l'ensemble de ces mesures de probabilité par :

$$\Pi(\mu, \nu) = \left\{ \pi \in \mathcal{M}_+^1(X \times Y) : (1.2) \text{ est vraie pour tous boréliens } A, B \right\},$$

où l'ensemble des mesures de probabilité sur $X \times Y$ est $\mathcal{M}_+^1(X \times Y)$. La version relaxée du problème (1.1) est alors

$$\inf \left\{ J(\pi) := \int_{X \times Y} c(x, y)\, d\pi(x, y) : \pi \in \Pi(\mu, \nu) \right\}. \tag{1.3}$$

Par ailleurs, nous pouvons écrire la condition (1.2) autrement :

$$\int_{X \times Y} (\varphi(x) + \psi(y))\, d\pi(x, y) = \int_X \varphi(x)\, d\mu(x) + \int_Y \psi(y)\, d\nu(y),$$

pour tout (φ, ψ) dans une classe adaptée de fonctions tests, par exemple $C_b(X) \times C_b(Y)$. Considérant cette réécriture, le problème dual est :

$$\sup \left\{ K(\varphi, \psi) := \int_X \varphi(x)\, d\mu(x) + \int_Y \psi(y)\, d\nu(y) : \varphi(x) + \psi(y) \leq c(x, y) \right\}. \tag{1.4}$$

Le problème (1.4) présente un point de vue assez différent. Alors que nous étions bloqués par une contrainte forte de non-linéarité dans le problème (1.1), nous avons maintenant besoin de trouver seulement une paire optimale (φ, ψ). En effet, la structure mathématique du problème (1.4) fournit précisément ce qui manquait dans le problème initial : une compacité assez bonne pour construire un élément minimisant comme sorte de limite d'une suite minimisante. C'est la "méthode directe en calcul des variations", qui est simplement le théorème classique de Weierstrass. Le théorème de dualité de Kantorovich dit que sous certaines hypothèses sur la structure des espaces mesurés et sur la fonction de coût, les problèmes (1.3) et (1.4) sont équivalents et que les deux ont des solutions :

$$\min_{\pi \in \Pi(\mu, \nu)} J(\pi) = \max\{K(\varphi, \psi) : \varphi(x) + \psi(y) \leq c(x, y)\}.$$

L'outil principal utilisé est le théorème de Rockafellar [54, 86], classique dans les problèmes de dualité en analyse convexe. En outre, nous pouvons caractériser ces solutions optimales. Pour plus de détails, nous renvoyons notamment à [9, 91, 94, 95]. Par ailleurs, la question de l'existence et la caractérisation géométrique des applications optimales T dans le problème (1.1) sont des champs de recherche qui se sont développés durant ces dernières vingt-cinq années. Pour la fonction de coût quadratique $c(x,y) = |x-y|^2/2$, nous pouvons consulter les notes de Evans [56]. Ce cas a été résolu par Brenier [36], avec son théorème de factorisation polaire qui dit que toute application "non-dégénérée" à valeurs vectorielles peut se réécrire comme la composée du gradient d'une fonction convexe et d'une application préservant la mesure de Lebesgue. Ambrosio et Pratelli [10] ont fait une revue très complète des résultats dans le cas original $c(x,y) = |x-y|$.

1.2 Problèmes de Beckmann

En parallèle des travaux de Kantorovich (dont il n'était pas au courant), dans un article d'économie, Beckmann [16] a proposé, dans les années cinquante, le problème suivant de minimisation de flux, appelé modèle continu de transport :

$$\min\left\{\int \mathcal{G}(\sigma)\, dx : \operatorname{div}\sigma = \mu - \nu\right\}, \tag{1.5}$$

pour $\mathcal{G}$ une fonction convexe et μ, ν deux mesures de probabilité sur Ω. C'est un problème particulier d'optimisation convexe. Le cas $\mathcal{G}(\sigma) = |\sigma|$ est intéressant parce qu'il permet de faire le lien avec le problème de Monge-Kantorovich (1.3), pour $c(x,y) = |x-y|$. Ce rapprochement a été fait par Feldman-McCann [58]. Le problème (1.5) devient :

$$\min\left\{\int_\Omega |\sigma(x)|\, dx : \sigma : \Omega \mapsto \mathbb{R}^d, \operatorname{div}\sigma = \mu - \nu\right\}, \tag{1.6}$$

où Ω est un domaine borné dans $\mathbb{R}^d$. Ici, la contrainte doit se comprendre au sens faible, avec les conditions de bord de Neumann, c'est-à-dire, pour toute fonction $u \in C^1(\overline{\Omega})$, nous avons

$$-\int_\Omega \nabla u \cdot \sigma = \int_\Omega u\, d(\mu - \nu).$$

Le problème dual est formellement :

$$\sup\left\{\int_\Omega \varphi\, d(\mu-\nu) : |\nabla\varphi| \leq 1\right\}.$$

C'est exactement le même problème que (1.4) car les fonctions 1-lipschitziennes sont les fonctions dont le gradient est plus petit que 1. On a ainsi une équivalence formelle entre les problèmes (1.6) et (1.3) ; c'est un cas particulier du théorème de Kantorovich-Rubinstein (pour une version générale, voir Dudley [51, 52]). Plus rigoureusement, si nous supposons que Ω est un domaine compact et convexe dans $\mathbb{R}^d$ alors le problème (1.6) admet une solution et sa valeur minimale est égale à celle du problème (1.3). De plus, nous pouvons construire une solution de (1.6) à partir d'une solution de (1.3). Une preuve se trouve dans [91]. L'inégalité $\min(1.3) \leq \min(1.6)$ est assez directe et pour l'inégalité inverse, nous prenons un plan de transport optimal γ pour (1.3). Nous construisons alors la mesure vectorielle σ^γ comme suit :

$$\int_\Omega \phi\, d\sigma^\gamma := \int_{\Omega\times\Omega}\int_0^1 (y-x)\cdot\phi((1-t)x+ty)\, dt\, d\gamma(x,y),$$

pour toute fonction $\phi \in C(\Omega, \mathbb{R}^d)$. De même, nous définissons la mesure scalaire m^γ ainsi :

$$\int_\Omega \xi \, dm^\gamma := \int_{\Omega\times\Omega} \int_0^1 |x-y|\xi((1-t)x+ty)\, dt\, d\gamma,$$

pour toute fonction $\xi \in C(\Omega, \mathbb{R})$. Cette mesure m^γ est appelée *densité de transport*. Nous trouvons alors que $\operatorname{div} \sigma^\gamma = \mu - \nu$ et que $|\sigma^\gamma| \leq m^\gamma$, où $|\sigma^\gamma|$ est la variation totale de la mesure vectorielle σ^γ. Finalement, nous avons $|\sigma^\gamma|(\Omega) \leq \min(1.3)$ et l'égalité $\min(1.3) = \min(1.6)$ est ainsi prouvée. Par abus de notations, on identifie une mesure σ avec la dérivée de Radon-Nikodym de σ par rapport à la mesure de Lebesgue. Evans-Gangbo [57] et Bouchitté-Buttazzo-Seppecher [26, 27] ont eu l'idée d'introduire les mesures m^γ et σ^γ pour des raisons différentes. Cette notion de densité de transport sera ensuite généralisée par Carlier-Jimenez-Santambrogio [39] aux mesures sur des courbes. Tout d'abord, soient un compact Ω d'intérieur non vide dans $\mathbb{R}^d$, une courbe absolument continue $\varphi : [0,1] \mapsto \Omega$ et une fonction continue $\xi : \Omega \mapsto \mathbb{R}_+$, définissons la longueur de φ pondérée avec le poids ξ :

$$L_\xi(\varphi) := \int_0^1 \xi(\varphi(t))|\varphi'(t)|\, dt. \tag{1.7}$$

C'est bien défini car les fonctions absolument continues sont différentiables presque partout et leur différentielle est dans $L^1([0,1])$. Nous notons l'ensemble des fonctions absolument continues sur $[0,1]$ et à valeurs dans Ω par C et nous le munissons de la topologie uniforme. Nous considérons alors une mesure de probabilité Q sur l'espace C telle que $\int_C L_1(\varphi)\, dQ(\varphi) < +\infty$. Une telle mesure sera appelée *plan de trafic* d'après la terminologie introduite par Bernot-Morel-Caselles [25]. Nous écrivons maintenant l'analogue de m^γ et σ^γ pour les mesures Q. Tout d'abord, l'*intensité du trafic*, notée par $i_Q \in \mathcal{M}_+(\Omega)$, est définie par :

$$\int_\Omega \xi \, di_Q := \int_C \left(\int_0^1 \xi(\varphi(t))|\varphi'(t)|\, dt\right) dQ(\varphi) = \int_C L_\xi(\varphi)\, dQ(\varphi), \tag{1.8}$$

pour tout $\xi \in C(\Omega, \mathbb{R}_+)$. C'est une généralisation de la notion de densité de transport. Pour un sous-ensemble borélien A de Ω, $i_Q(A)$ représente le trafic cumulé dans la zone A et induit par Q. De même, le *flux de trafic* induit par le plan de trafic Q est la mesure vectorielle σ^Q définie par :

$$\int_\Omega \phi \cdot d\sigma^Q := \int_C \left(\int_0^1 \phi(\varphi(t)) \cdot \varphi'(t)\, dt\right) dQ(\varphi), \tag{1.9}$$

pour tout $\phi \in C(\Omega, \mathbb{R}^d)$. Maintenant, nous nous intéressons seulement aux plans de trafic admissibles Q, c'est-à-dire, les plans de trafic Q tels que $e_{0\#}Q = \mu$ et $e_{1\#}Q = \nu$, e_0 et e_1 étant les évaluations en 0 et en 1. Nous notons $\mathcal{Q}(\mu,\nu)$ cet ensemble des plans de trafic admissibles. Alors nous avons facilement que $\operatorname{div} \sigma^Q = \mu - \nu$ et $|\sigma^Q| \leq i_Q$.

Réciproquement, Smirnov [92] et, plus tard, Santambrogio [90] ont prouvé que pour toute mesure vectorielle finie σ sur Ω et toutes mesures $\mu, \nu \in \mathcal{M}_+^1(\Omega)$ telles que $\operatorname{div} \sigma = \mu - \nu$, il existe un plan de trafic $Q \in \mathcal{Q}(\mu,\nu)$ tel que $|\sigma^Q| = i_Q \leq |\sigma|$ et

$$\|\sigma - \sigma^Q\| + \|\sigma^Q\| = \|\sigma - \sigma^Q\| + i_Q(\Omega) = \|\sigma\|,$$

où $\|\sigma\| = |\sigma|(\Omega)$. En particulier, si $\sigma^Q \neq \sigma$ alors $|\sigma^Q| \neq |\sigma|$.

Une généralisation possible du problème (1.5) avec une distance riemannienne $k(x)$ est

$$\min\left\{\int_\Omega k(x)|\sigma(x)|\,dx : \operatorname{div}\sigma = \mu-\nu\right\},$$

ce qui correspond, par dualité avec les functions u telles que $|\nabla u| \leq k$, à

$$\min\left\{\int_{\Omega\times\Omega} d_k(x,y)\,d\gamma(x,y) : \gamma\in\Pi(\mu,\nu)\right\},$$

où $d_k(x,y) := \inf\left\{\int_0^1 k(\varphi(t))|\varphi'(t)|\,dt : \varphi(0)=x, \varphi(1)=y\right\}$ est la distance associée à la métrique riemannienne k. La fonction $k(x)$ est le coût local en x par unité de longueur d'un chemin passant par x. La métrique est ainsi non-homogène. Ce modèle est plus pertinent si nous voulons prendre un coût non-uniforme pour le mouvement (dû aux obstacles géographiques ou aux configurations). Cependant, il n'est pas satisfaisant pour modéliser le trafic urbain congestionné. En effet, la métrique n'est à priori pas connue et dépend du trafic lui-même. Une variante possible, étudiée en particulier dans Carlier-Santambrogio [40], est de considérer les fonctions $k(x) = g(|\sigma(x)|)$, c'est-à-dire, k est une fonction du module du champ de vecteurs σ. Cela revient alors à résoudre le problème (1.5) avec $\mathcal{G}(z) = G(|z|)$ et $G(t) = g(t)t$. Dans le cas le plus simple $g(t) = t$, le problème consiste à minimiser la norme L^2 sous les contraintes de divergence :

$$\min\left\{\int_\Omega |\sigma(x)|^2 dx : \sigma\in L^2(\Omega,\mathbb{R}^d), \operatorname{div}\sigma = \mu-\nu\right\}.$$

Nous verrons plus tard que les problèmes à la Beckmann peuvent être reliés à des problèmes d'équilibre.

1.3 Équilibre de Wardrop discret

Un concept essentiel dans les problèmes de congestion est la notion de l'équilibre de Wardrop introduite dans les années cinquante par Wardrop [98]. Il présente de nombreuses applications dans des domaines variés : le trafic routier (Roughgarden-Tardos [88], Smith [93]), les jeux de congestion (Roughgarden-Tardos [87]), les réseaux de communication (Altman-Wynter [6]), les files d'attente (Altman-El Azouzi-Abramov [5]), les réseaux de transit (Cominetti-Correa [43]), etc. Nombre de ces articles utilisent également l'équilibre de Nash ([83]), en théorie de jeux, qui correspond au cas d'un nombre fini d'agents, qui décident d'une stratégie en sachant que leur utilité dépend des choix de tous.

Nous considérons d'abord un réseau discret. Pour modéliser le réseau, nous nous donnons donc un graphe fini orienté $G = (N, E)$, où N est l'ensemble des noeuds et E celui des arcs. Nous notons $t(e)$ le temps de parcours de l'arc e et $m(e)$ la masse se déplaçant sur l'arc e. Pour traduire les effets de la congestion, nous avons la relation :

$$t(e) = g(e, m(e))$$

où pour tout arc $e\in E$, la fonction $g(e,\cdot)$ est strictement positive, continue et croissante et prend en compte les effets de la congestion (qui dépendent de l'arc e selon ses caractéristiques : largeur, obstacles...). Nous écrivons $\mathbf{m}$ pour l'ensemble de toutes les masses sur les arcs $(m(e))_{e\in E}$. Une autre donnée du problème est un plan de transport γ sur les paires de noeuds $(x, y) \in N^2$. L'élément $\gamma(x, y)$ représente la masse qui doit être envoyée de la source x à la destination y. Un chemin générique φ est une suite finie de noeuds successifs et nous notons par $C_{x,y}$ l'ensemble des chemins allant de x à y, de sorte que

$$C := \bigcup_{(x,y)\in N^2} C_{x,y}$$

est l'ensemble de tous les chemins. Comme le temps de voyage sur chaque arc est strictement positif, nous pouvons considérer uniquement les chemins simples (sans boucle). Nous disons que $e \in \varphi$ si le chemin φ passe par l'arc e. La masse se déplaçant sur le chemin φ est notée $w(\varphi)$. La collection de toutes les masses sur les chemins $w(\varphi)$ est notée $\mathbf{w}$. Étant connues les masses sur les arcs $\mathbf{m}$, le temps de parcours d'un chemin $\varphi \in C$ est donné par :

$$\tau_{\mathbf{m}}(\varphi) = \sum_{e\in\varphi} g(e, m(e)).$$

En résumé, les inconnues du problème sont les collections de nombres positifs $\mathbf{m} = (m(e))_{e\in E}$ et $\mathbf{w} = (w(\varphi))_{\varphi\in C}$. Elles doivent vérifier les relations de conservation de la masse suivantes :

$$\gamma(x, y) = \sum_{\varphi\in C_{x,y}} w(\varphi), \quad \text{pour tout } (x, y) \in N \times N \tag{1.10}$$

et

$$m(e) = \sum_{\varphi\in C: e\in\varphi} w(\varphi), \quad \text{pour tout } e \in E. \tag{1.11}$$

Nous arrivons alors au concept d'équilibre de Wardrop. Ce concept exprime qu'à l'équilibre, seulement les chemins les plus courts (en prenant en compte la congestion créée par les masses sur les arcs et sur les chemins) sont utilisés. Nous supposons donc le conducteur rationnel : il prendra toujours un chemin optimal. Plus précisément, c'est la définition suivante :

Définition 1.1. *Un équilibre de Wardrop est une configuration de masses sur les arcs positives* $\mathbf{m} : e \mapsto m(e)$ *et de masses sur les chemins positives* $\mathbf{w} : \varphi \mapsto w(\varphi)$ *satisfaisant les relations de conservation de la masse* (1.10) *et* (1.11) *et telle que pour tout* $(x, y) \in N \times N$ *et tout* $\varphi \in C_{x,y}$*, si* $w(\varphi) > 0$ *alors nous avons :*

$$\tau_{\mathbf{m}}(\varphi) \leq \tau_{\mathbf{m}}(\varphi'), \quad \textit{pour tout } \varphi' \in C_{x,y}.$$

Quelques années plus tard, Beckmann, McGuire et Winsten [17] réalisèrent que l'équilibre de Wardrop pouvait être caractérisé par le principe variationnel suivant :

Théorème 1.1. *Une configuration de flux* $(\mathbf{w}, \mathbf{m})$ *est un équilibre de Wardrop si et seulement si elle minimise*

$$\sum_{e\in E} G(e, m(e)) \text{ où } G(e, m) := \int_0^m g(e, \alpha) d\alpha \tag{1.12}$$

sous les contraintes de positivité et aux relations de conservation de la masse (1.10)-(1.11).

Remarquons que le problème (1.12) est en fait un problème de minimisation uniquement en $\mathbf{w}$ car nous pouvons déduire $\mathbf{m}$ de $\mathbf{w}$ grâce à la condition (1.11). Comme les fonctions $g(e,\cdot)$ sont croissantes, le problème (1.12) est convexe donc nous avons facilement des résultats d'existence et nous pouvons approcher numériquement les solutions. Cependant, la complexité peut s'avérer très élevée si le réseau est très dense (comme c'est le cas notamment pour des réseaux réalistes de congestion). Donc nous pouvons préférer travailler avec le problème dual qui est le suivant :

$$\inf_{\mathbf{t}\in\mathbb{R}_+^{\#E}} \sum_{e\in E} H(e,t(e)) - \sum_{(x,y)\in N\times N} \gamma(x,y)T_{\mathbf{t}}(x,y), \tag{1.13}$$

où $\mathbf{t} \in \mathbb{R}_+^{\#E}$ signifie que $\mathbf{t} = (t(e))_{e\in E}$, $H(e,\cdot) := (G(e,\cdot))^*$ est la transformée de Legendre de $G(e,\cdot)$, qui est

$$H(e,t) := \sup_{m\geq 0}\{mt - G(e,m)\}, \quad \text{pour tout } t \in \mathbb{R}_+$$

et $T_{\mathbf{t}}$ est la fonctionnelle de longueur minimale :

$$T_{\mathbf{t}}(x,y) = \min_{\varphi\in C_{x,y}} \sum_{e\in\varphi} t(e).$$

Ici, nous avons besoin de connaître seulement $\#E$ variables donc la complexité semble meilleure. Cependant, un inconvénient majeur apparaît dans le problème (1.13) : c'est le terme qui dépend de $T_{\mathbf{t}}$. En effet, $T_{\mathbf{t}}$ est non-régulier, non-local et nous pouvons rencontrer des difficultés pour optimiser ce terme. Fukushima [60] a montré que si $(\mathbf{w},\mathbf{m})$ est un équilibre de Wardrop alors $\mathbf{t} := (g(e,m(e)))_{e\in E}$ est une solution de (1.13). Ainsi, résoudre le problème (1.13) revient à trouver les temps de parcours de l'équilibre et donc les masses sur les arcs $m(e)$ correspondantes en inversant la relation $t(e) = g(e,m(e))$. Une extension du modèle, au cas markovien (avec les temps qui sont inconnus), est également proposée dans ce papier.

Dans le problème présenté ci-dessus, le plan de transport γ est fixé. C'est le problème de court terme. Au lieu de cela, nous pourrions également considérer le cas où seulement ses marginales sont fixées. Plus précisément, nous nous donnons une distribution de sources $\mu = \sum_{x\in N}\mu(x)\delta_x$ et de destinations $\nu = \sum_{x\in N}\nu(x)\delta_x$ qui sont des mesures discrètes avec la même masse totale sur l'ensemble des noeuds N (que nous pouvons supposer égale à 1) :

$$\sum_{x\in N}\mu(x) = \sum_{x\in N}\nu(x) = 1.$$

Les nombres $\mu(x)$ et $\nu(x)$ sont positifs pour tout $x \in N$. Ceci peut être interprété comme un problème de long terme. Comme précédemment, nous pouvons reprendre presque les mêmes notations et la définition 1.1 de l'équilibre de Wardrop doit être légèrement modifiée. Nous devons remplacer la condition de conservation de la masse (1.10) par :

$$\mu(x) := \sum_{\varphi\in C_{x,\cdot}} w(\varphi),\ \nu(y) := \sum_{\varphi\in C_{\cdot,y}} w(\varphi) \tag{1.14}$$

pour tout $(x,y)\in N\times N$, où $C_{x,\cdot}$, respectivement $C_{\cdot,y}$, est l'ensemble des chemins simples commençant à l'origine x, respectivement finissant au point terminal y. En outre, le plan de transport est désormais une inconnue et nous devons ajouter une condition d'optimalité supplémentaire. Pour être plus précis, il existe un plan de transport optimal entre les marginales pour le coût de transport induit par la métrique congestionnée elle-même. Nous devons d'abord définir la fonctionnelle de longueur minimale $\tilde{T}_m$:

$$\tilde{T}_m(x,y) := \min_{\varphi \in C_{x,y}} \sum_{e\in\varphi} g(e, m(e)).$$

Soit $\Pi(\mu,\nu)$ l'ensemble des plans de transport discrets entre μ et ν, c'est-à-dire, l'ensemble de collections d'éléments positifs $(\gamma(x,y))_{(x,y)\in N\times N}$ telle que

$$\sum_{y\in N} \gamma(x,y) = \mu(x) \text{ et } \sum_{x\in N} \gamma(x,y) = \nu(y), \text{ pour tout } (x,y) \in N \times N.$$

Alors nous pouvons donner la définition de l'équilibre de Wardrop pour le modèle de long terme :

Définition 1.2. *Un équilibre de Wardrop est une configuration de masses sur les arcs positives* $\mathbf{m} : e \mapsto m(e)$ *et de masses sur les chemins positives* $\mathbf{w} : \varphi \mapsto w(\varphi)$ *satisfaisant les relations de conservation de la masse* (1.14) *et* (1.11) *et telle que :*

1. *Pour tout* $(x,y) \in N \times N$ *et tout* $\varphi \in C_{x,y}$, *si* $w(\varphi) > 0$ *alors nous avons :*

$$\tau_{\mathbf{m}}(\varphi) = \min_{\varphi' \in C_{x,y}} \tau_{\mathbf{m}}(\varphi'), \tag{1.15}$$

2. *Si nous définissons* $f(x,y) = \sum_{\varphi\in C_{x,y}} w(\varphi)$ *alors* f *est un élément minimisant de*

$$\inf_{\gamma\in\Pi(\mu,\nu)} \sum_{(x,y)\in N\times N} \gamma(x,y)\tilde{T}_m(x,y). \tag{1.16}$$

La deuxième condition est un problème de transport discrétisé de (1.3). Des arguments similaires à ceux dans le cas de court terme s'appliquent ici, l'équilibre est un élément minimisant de la fonctionnelle définie par (1.12) mais maintenant sous les contraintes de positivité et les conditions (1.14) et (1.11). L'analogue de la formulation duale (1.13) est alors :

$$\inf_{\mathbf{t}\in\mathbb{R}_+^{\#E}} \left\{ \sum_{e\in E} H(e,t(e)) - \inf_{\gamma\in\Pi(\mu,\nu)} \sum_{(x,y)\in N\times N} \gamma(x,y) T_{\mathbf{t}}(x,y) \right\}. \tag{1.17}$$

Le modèle de long terme a été étudié en particulier dans [31, 33, 66] tandis que dans [12, 67] notamment, les auteurs ont travaillé avec une variante de court terme.

1.4 Équilibre de Wardrop continu

La sous-section précédente présentait l'équilibre de Wardrop dans un cadre discret. Nous pouvons généraliser cette notion à un équilibre continu. Pour cela, nous allons utiliser des mesures de probabilité Q sur l'ensemble des chemins pour traduire la dépendance du modèle de transport par rapport aux chemins. Ces mesures Q sont les analogues continues des flux de chemin $(w(\varphi))_\varphi$. De même, nous mesurerons l'intensité du trafic généré par Q en chaque point x grâce à la mesure i_Q (définie par (1.8)), qui est l'équivalent continu des flux d'arc $(m(e))_e$. La dernière donnée est une métrique, qui est croissante par rapport à l'intensité du trafic, pour modéliser les effets de la congestion (elle est comparable à $g(e, i(e))$). Prenons Ω un domaine de $\mathbb{R}^d$ et un ensemble convexe et fermé $\Gamma \subset \Pi(\mu, \nu)$, où μ et ν sont des mesures de probabilité sur Ω. Le modèle de court terme est le cas particulier $\Gamma = \{\gamma\}$ et pour se ramener au cas de long terme, il suffit de prendre $\Gamma = \Pi(\mu, \nu)$.

Plus précisément, nous considérons des mesures de probabilité Q sur l'ensemble C des fonctions définies sur $[0, 1]$, absolument continues, à valeurs dans Ω et compatibles avec la conservation de la masse, c'est-à-dire, telles que $(e_0, e_1)_\# Q \in \Gamma$. L'ensemble de telles mesures est noté $\mathcal{Q}(\Gamma)$. Nous supposons que i_Q est absolument continue par rapport à $\mathcal{L}^d$ et la métrique est

$$\xi_Q(x) := g(x, i_Q(x))$$

avec $g : \Omega \times \mathbb{R}_+ \mapsto \mathbb{R}_+$ croissante par rapport à la seconde variable. Plutôt que $i_Q \ll \mathcal{L}(\mathbb{R}^d)$, nous pouvons également supposer $i_Q \in L^q(\Omega)$ avec $1 \leq q \leq +\infty$. La question de l'existence d'une telle mesure Q n'est pas triviale ; elle dépend des hypothèses faites sur μ, ν, Γ, q et Ω. Elle a été traitée notamment par Benmansour-Carlier-Peyré-Santambrogio [23] dans le cas de court terme (avec γ étant une mesure discrète de probabilité, et $q \in [1, 2)$) et par De Pascale-Evans-Pratelli [47–49] et Santambrogio [89] dans le cas de long-terme (si Ω est convexe et μ et ν sont dans $L^q(\Omega)$). Pour des problématiques liées à la mécanique des fluides incompressibles, Brenier [35] a construit un Q tel que $i_Q \in L^\infty$ si μ et ν sont dans L^∞, dans le cas $\Gamma = \{\gamma\}$. Brasco-Petrache [34] a montré qu'une condition nécessaire et suffisante pour avoir l'existence d'un Q tel que $i_Q \in L^q(\Omega)$ $(1 \leq q \leq \infty)$ est que $\mu - \nu \in W^{-1,q}(\Omega)$.

Afin de définir la notion d'équilibre de Wardrop dans le cas continu, définissons d'abord

$$L_{\xi_Q}(\varphi) := \int_0^1 \xi_Q(\varphi(t))|\varphi'(t)|dt = \int_0^1 g(\varphi(t), i_Q(\varphi(t)))|\varphi'(t)|dt \tag{1.18}$$

et

$$c_{\xi_Q}(x, y) := \inf \left\{ L_{\xi_Q}(\varphi) : \varphi \in C, \varphi(0) = x, \varphi(1) = y \right\}. \tag{1.19}$$

Les chemins dans C tels que $c_{\xi_Q}(\varphi(0), \varphi(1)) = L_{\xi_Q}(\varphi)$ sont appelés géodésiques (pour la métrique induite par les effets de la congestion générée par Q).

Alors nous pouvons écrire la définition de l'équilibre de Wardrop continu :

Définition 1.3. *Un équilibre de Wardrop est une mesure $Q \in \mathcal{Q}(\Gamma)$ telle que*

$$Q(\{\varphi : L_{\xi_Q}(\varphi) = c_\xi(\varphi(0), \varphi(1))\}) = 1.$$

Cependant, la définition (1.18) de $L_\xi(\varphi)$ n'a de sens que pour ξ continue et $\varphi \in C$. La généralisation au cas ξ mesurable et φ lipschitzienne n'est pas simple. Carlier-Jimenez-Santambrogio [39] propose une contsruction de c_ξ quand ξ est seulement L^p, à valeurs positives, avec $1 < p < +\infty$. Cette construction est en particulier utilisée dans Baillon-Carlier [12] et Hatchi [67]. Ainsi, l'existence d'un tel équilibre n'est pas triviale. Néanmoins, toujours dans [39], il est montré que l'équilibre pouvait être vu comme une solution d'un problème variationnel, qui est le suivant :

$$\min \left\{ \int_\Omega G(x, i_Q(x))\, dx : Q \in \mathcal{Q}(\Gamma) \right\}, \tag{1.20}$$

où $G(x, m) = \int_0^m g(x, \alpha)\, d\alpha$. C'est la version continue de (1.12). Le résultat principal de [39] (sous certaines hypothèses techniques) est alors :

Théorème 1.2. *Le problème* (1.20) *admet au moins un élément minimisant. De plus, $\overline{Q} \in \mathcal{Q}(\Gamma)$ est une solution si et seulement si c'est un équilibre de Wardrop et $\gamma_{\overline{Q}} := (e_0, e_1)_\# \overline{Q}$ est une solution du problème d'optimisation*

$$\min \left\{ \int_{\Omega\times\Omega} c_{\xi_Q}(x,y)\, d\gamma(x,y) : \gamma \in \Gamma \right\}.$$

En particulier, dans le cas de court terme ($\Gamma = \{\gamma\}$), la dernière condition devient triviale et nous avons l'existence d'un équilibre de Wardrop correspondant à un plan de transport γ donné. Dans le cas de long terme ($\Gamma = \Pi(\mu,\nu)$), cela signifie que γ_Q est une solution du problème de Monge-Kantorovich pour un coût de distance qui dépend de Q lui-même et on a ainsi une nouvelle condition d'équilibre. Ce résultat est généralisé dans [12, 67].

On constate ainsi que les équilibres de Wardrop admettent une formulation variationnelle, qui est en principe plus simple à étudier que la définition elle-même, comme dans le cas d'un réseau discret. Cependant, le problème (1.20) implique des mesures sur des ensembles de courbes ; il peut donc être délicat de le résoudre. En effet, nous avons deux couches de dimensions infinies. Néanmoins, dans le cas de long terme, nous allons voir dans la sous-section suivante que le problème (1.20) peut se reformuler en un problème à la Beckmann.

Remarquons aussi que le problème (1.20) est l'analogue de (1.12). La version continue de (1.13)-(1.17) est :

$$\inf_{\xi\geq 0} \left\{ \int_\Omega H(x,\xi(x))\, dx - \inf_{\gamma\in\Pi(\mu,\nu)} \int_{\Omega\times\Omega} c_\xi(x,y)\, d\gamma(x,y) \right\}. \tag{1.21}$$

Dans le cas de court terme ($\Gamma = \{\gamma\}$), nous avons un résultat de dualité [23] :

Proposition 1.1. *Si $\Gamma = \{\gamma\}$ et $\mathcal{Q}(\gamma) \neq \emptyset$, nous avons*

$$\min (1.20) = \max (1.21).$$

De plus, ξ est une solution de (1.21) *si et seulement si $\xi = \xi_Q$ pour un $Q \in \mathcal{Q}(\gamma)$ solution de* (1.20).

A partir de ce résultat, les mêmes auteurs [23, 24] ont proposé une méthode numérique consistante pour approcher la métrique ξ_Q solution du problème dual. Elle se base sur le *Fast Marching Method*. Elle consiste à travailler avec une version discrète du problème dual et à utiliser une méthode de descente de gradient.

1.5 Équivalence avec un problème à la Beckmann et EDPs dans le cas de long terme

Dans cette sous-section, nous fixons $\Gamma = \Pi(\mu,\nu)$, nous sommes dans le cadre de long terme. Alors dans le problème (1.20), comme le plan de transport n'est pas fixé, nous avons un degré de liberté supplémentaire. Cela permettra de reformuler (1.20) en tant que problème à la Beckmann avec une contrainte sur la divergence et ainsi de réduire le problème d'équilibre à la résolution d'une EDP non-linéaire. Rappelant la définition (1.9) de la mesure vectorielle σ^Q pour $Q \in \mathcal{Q}(\Gamma)$, il est immédiat de voir que la valeur de (1.20) est supérieure à celle du problème de flux minimal à la Beckmann suivant :

$$\min \left\{ \int_\Omega \mathcal{G}(x, \sigma(x))\, dx : \operatorname{div} \sigma = \mu - \nu \right\}, \tag{1.22}$$

où $\mathcal{G}(x, \sigma) = G(x, |\sigma|)$. En effet, nous utilisons l'inégalité $|\sigma^Q| \le i_Q$ et le fait que $G(x, \cdot)$ est croissante. Pour prouver l'inégalité inverse, une idée naturelle est de construire, à partir d'un σ admissible, un $Q \in \mathcal{Q}(\Gamma)$ tel que $i_Q = |\sigma|$.

En omettant les problèmes de régularité, un candidat Q est donné par

$$Q := \int_\Omega \delta_{X.(x)} d\mu_0(x)$$

où X est le flot du champ de vecteurs non-autonome $v = \sigma/f$:

$$\begin{cases} \partial_t X_t(x) = v(t, X_t(x)), \\ \quad X_0(x) = x, \ (t, x) \in [0, 1] \times \Omega, \end{cases}$$

avec $f(t, x) := (1-t)\mu(x) + t\nu(x)$ pour tous $t \in [0, 1]$ et $x \in \Omega$. En effet, f est solution de l'équation de continuité $\partial_t f + \operatorname{div}(fv) = 0$ avec la donnée initiale $f(0, \cdot) = \mu$. Si v est assez régulière par rapport à x (lipschitzienne par exemple), le problème de Cauchy pour l'équation de continuité

$$\begin{cases} \partial_t \varrho + \operatorname{div}(\varrho v) = 0, \\ \qquad \varrho(0, \cdot) = \mu(x), \end{cases}$$

a une solution unique pour chaque donnée initiale x qui est $X._\# \mu$. C'est un cas particulier de la méthode des caractéristiques (voir [8] pour la théorie). Par ailleurs, f vérifie le même problème de Cauchy donc $f = X._\# \mu$. Cette construction repose sur l'argument de flot de Dacorogna et Moser [44, 82], utilisé pour la première fois en transport optimal par Evans-Gangbo [57]. Cet argument est repris pour des problèmes de Beckmann par Brasco, Carlier et Santambrogio [31, 33] pour montrer l'égalité $\min (1.20) = \min (1.22)$. Cette égalité est intéressante. En effet, si on restreint le problème (1.22) à $\sigma \in L^q(\Omega, \mathbb{R}^d)$ $(q \in (1, +\infty))$, le problème dual de (1.22) est

$$\sup_{u \in W^{1,p}(\Omega)} \left\{ \int_\Omega u\, df - \int_\Omega \mathcal{G}^*(x, \nabla u(x))\, dx \right\}, \tag{1.23}$$

où p est l'exposant conjugué de q et $\mathcal{G}^*$ est la transformée de Legendre de $\mathcal{G}(x, \cdot)$. En utilisant des arguments classiques de dualité en analyse convexe ([54] par exemple), nous avons $\min (1.22) = \max (1.23)$ et les conditions d'optimalité primal-dual caractérisent l'élément minimisant σ (qui est unique si $\mathcal{G}(x, \cdot)$ est strictement convexe) de (1.22) par

$$\sigma(x) = \nabla \mathcal{G}^*(x, \nabla u(x)), \ \text{ pour presque tout } x \in \Omega, \tag{1.24}$$

où u est une solution de (1.23). Ceci est équivalent à la condition nécessaire que u est une solution faible de l'équation de Euler-Lagrange :

$$\begin{cases} \operatorname{div} \nabla \mathcal{G}^*(x, \nabla u) = \mu - \nu \ \text{ dans } \Omega, \\ \nabla \mathcal{G}^*(x, \nabla u) \cdot \nu_\Omega = 0 \ \text{ sur } \partial\Omega, \end{cases}$$

au sens que

$$\int_{\Omega} \nabla \mathcal{G}^*(x, \nabla u(x)) \cdot \nabla \varphi(x)\, dx = \int_{\overline{\Omega}} \varphi(x)\, d(\mu - \nu)(x), \quad \text{pour tout } \varphi \in W^{1,p}(\Omega).$$

Comme nous le verrons par la suite, cette reformulation permettra d'utiliser une méthode numérique pour approcher les solutions.

2 Contributions

2.1 Équilibre de Wardrop : étude rigoureuse de limites continues de modèles généraux de réseaux

Dans la première partie, nous avons vu les définitions de l'équilibre de Wardrop pour un réseau discret et également dans un cadre continu. Les questions qui se posent alors sont : que se passe-t-il quand un réseau discret devient très dense ? Est-ce que l'on a une convergence des valeurs et solutions des problèmes de minimisation discrets vers celles des problèmes continus ? Baillon-Carlier [12] a étudié le cas d'une grille cartésienne. Soient Ω un domaine borné de $\mathbb{R}^2$ avec une frontière régulière et $\varepsilon > 0$, ils ont considéré ce réseau discret (dont la longueur caractéristique est ε) :

$$\Omega_\varepsilon := \varepsilon\mathbb{Z}^2 \cap \overline{\Omega}.$$

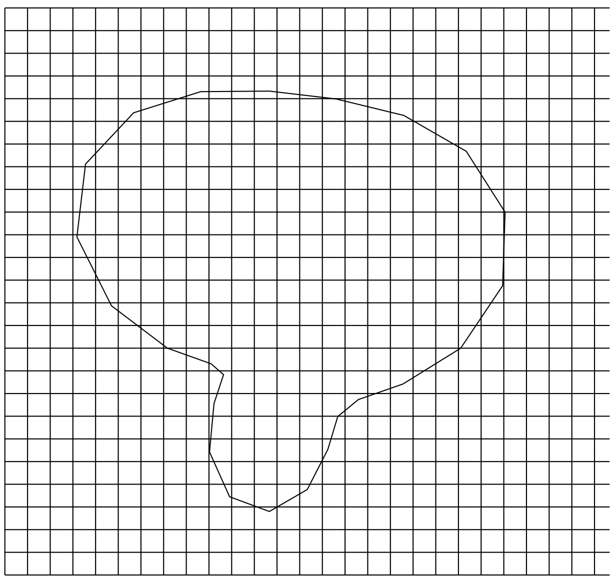

FIGURE 1.1 – Un réseau cartésien dans $\mathbb{R}^2$

Ici, les directions possibles sont $(v_1, v_2, v_3, v_4) := ((1,0), (0,1), (-1,0), (0,-1))$ en chaque noeud $x \in \Omega_\varepsilon$ du réseau et un arc est de la forme $[x, x + \varepsilon v_i]$ pour $x \in \Omega_\varepsilon$ et $i \in \{1, \dots, 4\}$. On identifiera ces arcs aux paires (x, v_i). Nous sommes dans un réseau discret donc nous reprenons les notations de la sous-section 1.3 en ajoutant l'exposant ε, pour traduire la dépendance en ε. Dans un premier temps, on considère le modèle de court terme avec un plan de transport γ^ε fixé. Le problème (1.12) devient alors :

$$\min_{(\mathbf{w}^\varepsilon,\mathbf{m}^\varepsilon)} \left\{ \sum_{x\in\Omega_\varepsilon} \sum_{i=1}^{4} G_i^\varepsilon(x, m_i^\varepsilon(x)) : (\mathbf{w}^\varepsilon, \mathbf{m}^\varepsilon) \geq 0 \text{ et (1.10)-(1.11) sont vérifiées} \right\} \tag{1.25}$$

où $G_i^\varepsilon(x,m) := \int_0^m g_i^\varepsilon(x,\alpha)d\alpha$, tandis que l'analogue du problème (1.13) est

$$\inf_{\mathbf{t}^\varepsilon\in\mathbb{R}_+^{4\#\Omega_\varepsilon}} \left\{ \sum_{x\in\Omega_\varepsilon} \sum_{i=1}^{4} H_i^\varepsilon(x, t_i^\varepsilon(x)) - \sum_{(x,y)\in\Omega_\varepsilon^2} \gamma^\varepsilon(x,y) T_{\mathbf{t}^\varepsilon}^\varepsilon(x,y) \right\}, \tag{1.26}$$

L'indice i signifie que la direction considérée est v_i : ainsi, $m_i^\varepsilon(x) = m^\varepsilon(x, v_i)$ représente la masse se déplaçant sur l'arc (x, v_i). Pour pouvoir passer à la limite quand ε tend vers 0^+, Baillon et Carlier [12] ont fait des hypothèses. La première porte sur la famille des plans de transports $(\gamma^\varepsilon)_{\varepsilon>0}$: il existe une mesure finie et positive γ sur $\overline{\Omega}\times\overline{\Omega}$ vers laquelle γ^ε converge faiblement-$\star$:

$$\lim_{\varepsilon\to 0^+} \sum_{(x,y)\in\Omega_\varepsilon^2} \gamma^\varepsilon(x,y)\varphi(x,y) = \int_{\overline{\Omega}\times\overline{\Omega}} \varphi\, d\gamma, \ \text{ pour tout } \varphi \in C(\overline{\Omega}\times\overline{\Omega}). \tag{1.27}$$

La seconde hypothèse est sur les fonctions de congestion g^ε, qui sont supposées être de la forme suivante :

$$g_i^\varepsilon(x,m) = \varepsilon g_i\left(x, \frac{m}{\varepsilon}\right), \ \text{ pour tout } (x,\varepsilon,i) \in \Omega_\varepsilon \times \mathbb{R}_+^* \times \{1,\cdots,4\}, \tag{1.28}$$

où g_i est une fonction continue et positive sur $\overline{\Omega}\times\mathbb{R}_+$, qui est croissante par rapport à la seconde variable. Cette hypothèse signifie que le temps de parcours d'un arc de longueur ε est d'ordre ε et dépend du flux par unité de longueur m/ε. Cela nous permet de s'assurer de la stricte convexité des fonctions H^ε et H. En définissant la variable métrique (ou le temps par unité de longueur) $\xi^\varepsilon := \mathbf{t}^\varepsilon/\varepsilon$, on peut réécrire le problème (1.26) ainsi :

$$\inf_{\xi^\varepsilon\in\mathbb{R}_+^{4\#\Omega_\varepsilon}} J^\varepsilon(\xi^\varepsilon) := I_0^\varepsilon(\xi^\varepsilon) - I_1^\varepsilon(\xi^\varepsilon) \tag{1.29}$$

avec

$$I_0^\varepsilon(\xi^\varepsilon) := \varepsilon^2 \sum_{x\in\Omega_\varepsilon} \sum_{i=1}^{4} H_i(x, \xi_i^\varepsilon(x)) \tag{1.30}$$

et

$$I_1^\varepsilon(\xi^\varepsilon) := \varepsilon \sum_{(x,y)\in\Omega_\varepsilon^2} \gamma^\varepsilon(x,y) T_{\xi^\varepsilon}^\varepsilon(x,y). \tag{1.31}$$

La dernière hypothèse est que H_i est continu par rapport à la première variable et qu'il existe $p > 2$ et deux constantes $0 < \lambda < \Lambda$ telles que pour tout $(x,\xi,i) \in \Omega_\varepsilon \times \mathbb{R}_+ \times \{1,\dots,4\}$, nous avons :

$$\lambda(\xi^p - 1) \leq H_i(x,\xi) \leq \Lambda(\xi^p + 1). \tag{1.32}$$

Cette condition de croissance permet de travailler dans L^p à la limite et ainsi de construire des termes d'intégrales comme limite continue.

Compte-tenu des hypothèses faites, définissons

$$L^p_+ := \{\xi = (\xi_1, \ldots, \xi_4), \xi_i \in L^p(\Omega), \xi_i \geq 0, i = 1, \ldots, 4\}.$$

Comme limite continue du terme I_0^ε, il est naturel de penser à cette fonctionnelle intégrale

$$I_0(\xi) := \sum_{i=1}^{4} \int_\Omega H_i(x, \xi_i(x))\, dx, \ \text{ pour tout } \xi \in L^p_+. \tag{1.33}$$

Pour le terme I_1^ε, il est plus délicat de trouver une limite continue, à cause de la présence d'un terme non local et non régulier. Nous devons généraliser la définition de c_ξ donnée par (1.19). Pour $\xi \in C(\overline{\Omega}, \mathbb{R}^4_+), c_\xi$ est donné par la formule suivante :

$$c_\xi(x, y) := \inf \left\{ \sum_{i=1}^{4} \int_0^1 \xi_i(\sigma(t))(\dot\sigma(t) \cdot v_i)_+ \, dt \right\}, \ \text{ pour tous } x, y \in \overline{\Omega}, \tag{1.34}$$

où l'infimum est pris sur l'ensemble des courbes absolument continues σ, à valeurs dans $\overline{\Omega}$ et telles que $\sigma(0) = x$ et $\sigma(1) = y$. Ainsi, c_ξ rend compte de l'anisotropie du réseau. Pour étendre la définition de c_ξ à $\xi \in L^p_+$, nous généralisons la construction de $\overline{c}_\xi$ faite dans [39]. Alors la limite continue de J^ε est :

$$\inf_{\xi \in L^p_+} J(\xi) := I_0(\xi) - I_1(\xi) = \sum_{i=1}^{4} \int_\Omega H_i(x, \xi_i(x))\, dx - \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_\xi \, d\gamma. \tag{1.35}$$

C'est l'analogue de (1.21).

Pour montrer que (1.35) est bien la limite continue de (1.29), nous utilisons la théorie de Γ-convergence. C'est un outil puissant pour décrire le comportement asymptotique de familles de problèmes de minimisation habituellement dépendant de paramètres (d'échelle, géométriques, etc.). Depuis son introduction par De Giorgi (notamment [45, 46]) dans les années soixante-dix, c'est la notion de convergence la plus flexible et naturelle pour les problèmes variationnels. On a la convergence non seulement des valeurs mais également des éléments minimisants. Ici, notre problème a une dépendance en ε (qui est l'échelle du réseau Ω_ε) et nous voulons passer à la limite quand ε tend vers 0^+. Par conséquent, cette théorie est particulièrement indiquée. Pour la théorie générale de Γ-convergence, des références possibles sont les livres de Braides [29] et Dal Maso [80].

Nous voulons montrer la convergence d'une famille de fonctionnelles dans un cadre discret vers une fonctionnelle continue donc nous devons en particulier préciser dans quel sens la convergence d'une famille discrète $\xi^\varepsilon \in \mathbb{R}_+^{4\#\Omega_\varepsilon}$ vers un $\xi \in L^p_+$ se fait. Nous renvoyons au chapitre 2 pour les définitions de la topologie de L^p_+ et de la Γ-convergence (dans un cadre plus général que le modèle cartésien). Le principal résultat de [12] est alors :

Théorème 2.1. *Sous les hypothèses* (1.27), (1.28), (1.32), *la famille de fonctionnelles J^ε définies par* (1.29) Γ*-converge (pour la topologie faible de L^p) vers la fonctionnelle J définie par* (1.35).

Par des arguments classiques de la théorie générale de Γ-convergence, nous avons le résultat de convergence suivant :

Corollaire 2.1. *Sous les hypothèses* (1.27), (1.28), (1.32), *les problèmes* (1.29) *et* (1.35) *admettent une solution optimale et nous avons*

$$\min_{\xi^\varepsilon \in \mathbb{R}_+^{4\#\Omega_\varepsilon}} J^\varepsilon(\xi^\varepsilon) \to \min_{\xi \in L_+^p} J(\xi) \quad \textit{quand } \varepsilon \to 0^+.$$

De plus, pour tout $\varepsilon > 0$, si ξ^ε est l'élément minimisant de (1.29), *alors $\xi^\varepsilon \to \xi$, où ξ est l'élément minimisant de J sur L_+^p.*

C'est sous-entendu dans le corollaire, nous avons l'unicité de la solution ξ^ε, respectivement ξ, de (1.29), respectivement de (1.35). Cela vient de la stricte convexité de ces deux problèmes (les termes I_0^ε et I_0 sont strictement convexes tandis que les termes I_1^ε et I_1 sont concaves). Dans [67], j'ai généralisé les résultats de [12]. Au lieu de considérer une grille cartésienne dans $\mathbb{R}^2$ de plus en plus dense, j'ai étudié des réseaux discrets généraux dans $\mathbb{R}^d$ dont les directions et les longueurs d'arc ne sont pas fixées. Soit Ω un domaine borné de $\mathbb{R}^d$, nous analysons une suite de réseaux discrets $\Omega_\varepsilon = (N^\varepsilon, E^\varepsilon)$, où N^ε est l'ensemble des noeuds et E^ε celui des arcs. Un arc dans E^ε est de la forme (x, e) où $x \in N^\varepsilon$ et $e \in \mathbb{R}^d$ tels que le segment $[x, x+e]$ est inclus dans Ω. La longueur des arcs de Ω_ε est de l'ordre de $\varepsilon > 0$. Nous devons faire des hypothèses structurelles sur Ω_ε pour pouvoir passer à la limite. Pour cela, j'ai d'abord travaillé avec le cas $\Omega \subset \mathbb{R}^2$ discrétisé en hexagones réguliers. Ainsi, en chaque noeud dans N^ε, il y a 3 directions (constantes) possibles : (v_1, v_3, v_5) ou (v_2, v_4, v_6), où $v_k = \exp(i(\pi/6 + (k-1)\pi/3)), k = 1, \ldots, 6$. La principale différence par rapport au modèle cartésien est la perte de l'unicité de la décomposition conique (c'est-à-dire, à coefficients positifs) de tout $z \in \mathbb{R}^2$ dans l'ensemble de ces directions. Ensuite, j'ai construit des réseaux Ω_ε généraux dans $\Omega \subset \mathbb{R}^d$, où les directions dépendent du noeud et les arcs dans E^ε ne sont pas tous de même longueur. L'une des principales hypothèses faites est la suivante :

Hypothèse 2.1. *Il existe $N \in \mathbb{N}, D = \{v_k\}_{k=1,\ldots,N} \in C^{0,\alpha}(\mathbb{R}^d, \mathbb{S}^{d-1})^N$ et $\{c_k\}_{k=1,\ldots,N} \in C^1(\overline{\Omega}, \mathbb{R}_+^*)^N$ avec $\alpha > d/p$ tels que E^ε converge faiblement, au sens que*

$$\lim_{\varepsilon \to 0^+} \sum_{(x,e) \in E^\varepsilon} |e|^d \varphi\left(x, \frac{e}{|e|}\right) = \int_{\Omega \times \mathbb{S}^{d-1}} \varphi(x, v)\, \theta(dx, dv), \forall \varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}), \quad (1.36)$$

où $\theta \in \mathcal{M}_+(\Omega \times \mathbb{S}^{d-1})$ est de la forme

$$\theta(dx, dv) = \sum_{k=1}^{N} c_k(x) \delta_{v_k(x)} dx.$$

Les v_k représentent les directions dans le modèle continu et ne sont à priori pas constants. Les c_k sont les coefficients de volume. C'est une généralisation substantielle du cas cartésien, dans lequel θ est comme suit

$$\theta(dx, dv) = \left(\delta_{(1,0)} + \delta_{(0,1)} + \delta_{(-1,0)} + \delta_{(0,-1)}\right) dx.$$

Deux difficultés majeures apparaissent par rapport au modèle initial. Dans le modèle continu, nous avons N directions dans $\mathbb{S}^{d-1}$ et leur enveloppe conique est $\mathbb{R}^d$ entier. Mais pour tout $z \in \mathbb{R}^d$, nous avons à priori plusieurs décompositions coniques (non triviales) possibles. Dans le cas cartésien, nous avons une unique décomposition conique (non triviale) $z = \sum_{i=1}^{4}(z \cdot v_i)_+ v_i$, nous l'utilisons en particulier dans la définition (1.34) de c_ξ. Dans le cas général, nous devons ajouter un autre infimum sur les décompositions coniques possibles. Plus précisément, pour $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, c_ξ devient :

$$c_\xi(x,y) := \inf_{\sigma \in C_{x,y}} \int_0^1 \Phi_\xi(\sigma(t), \dot{\sigma}(t))\, dt \text{ pour tous } x, y \in \overline{\Omega}$$

avec pour tous $x \in \overline{\Omega}$ et $y \in \mathbb{R}^d$, $\Phi_\xi(x,y)$ étant défini comme suit :

$$\Phi_\xi(x,y) := \inf_{(y_1,\dots,y_N) \in \mathbb{R}_+^N} \left\{ \sum_{k=1}^{N} y_k \xi(x, v_k(x)) : y = \sum_{k=1}^{N} y_k v_k(x) \right\}.$$

Les $(y_1, \dots, y_N)$ sont les décompositions de y dans $(v_1(x), \dots, v_N(x))$. Comme précédemment, pour étendre cette définition à $\xi \in L^p_+(\theta) = \{\varphi \in L^p(\Omega \times \mathbb{S}^{d-1}, \theta), \varphi \geq 0\}$, nous généralisons les résultats de [39].

En outre, nous adaptons également les hypothèses (1.27), (1.28), (1.32). En particulier, (1.28) devient

Hypothèse 2.2. *g^ε est de la forme*

$$g^\varepsilon(x,e,m) = |e|^{d/2} g\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d/2}}\right), \forall \varepsilon > 0, (x,e) \in E^\varepsilon, m \geq 0 \tag{1.37}$$

où $g : \Omega \times \mathbb{S}^{d-1} \times \mathbb{R}_+ \mapsto \mathbb{R}$ est une fonction donnée, qui est continue, positive et croissante par rapport à la dernière variable.

Nous posons

$$\xi^\varepsilon(x,e) = \frac{t^\varepsilon(x,e)}{|e|^{d/2}} \text{ pour tout } (x,e) \in E^\varepsilon. \tag{1.38}$$

Si $d = 2$, (1.37) signifie que le temps de parcours d'un arc de longueur $|e|$ est de l'ordre de $|e|$ et dépend du flot par unité de longueur $m/|e|$. C'est naturel en termes de scaling. Nous avons défini dans (1.38) des variables métriques (c'est le temps par unité de longueur). Pour $d \neq 2$, le terme $d/2$ n'est pas très naturel mais il permet d'avoir le même exposant dans les définitions (1.37) et (1.38). L'analogue de (1.29) est alors :

$$\inf_{\xi^\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}} J^\varepsilon(\xi^\varepsilon) := I_0^\varepsilon(\xi^\varepsilon) - I_1^\varepsilon(\xi^\varepsilon) \tag{1.39}$$

où

$$I_0^\varepsilon(\xi^\varepsilon) := \sum_{(x,e) \in E^\varepsilon} |e|^d H\left(x, \frac{e}{|e|}, \xi^\varepsilon(x,e)\right) \tag{1.40}$$

et

$$I_1^\varepsilon(\xi^\varepsilon) := \sum_{(x,y) \in N^{\varepsilon 2}} \gamma^\varepsilon(x,y) \left(\min_{\sigma \in C^\varepsilon_{x,y}} \sum_{(z,e) \subset \sigma} |e|^{d/2} \xi^\varepsilon(z,e) \right).$$

La limite continue ressemble à celle dans le cas cartésien :

$$\inf_{\xi \in L^p_+(\theta)} J(\xi) := I_0(\xi) - I_1(\xi) = \int_{\Omega \times \mathbb{S}^{d-1}} H(x, v, \xi(x,v)) \theta(dx, dv) - \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_\xi d\gamma. \tag{1.41}$$

Dans ce chapitre, nous montrons qu'on arrive aux mêmes résultats de convergence que dans le cas cartésien.

Remarque 1. *(Analyse dimensionnelle) De manière plus générale, supposons au lieu de* (1.27)*,* (1.36)*,* (1.37) *et* (1.38) *que :*

$$\lim_{\varepsilon\to 0^+} \varepsilon^{\alpha_1} \sum_{(x,y)\in\Omega_\varepsilon^2} \gamma^\varepsilon(x,y)\varphi(x,y) = \int_{\overline{\Omega}\times\overline{\Omega}} \varphi \, d\gamma, \quad \textit{pour tout } \varphi \in C(\overline{\Omega}\times\overline{\Omega}),$$

$$\lim_{\varepsilon\to 0^+} \sum_{(x,e)\in E^\varepsilon} |e|^{\alpha_2} \varphi\left(x, \frac{e}{|e|}\right) = \int_{\overline{\Omega}\times\mathbb{S}^{d-1}} \varphi(x,v)\, \theta(dx,dv), \quad \textit{pour tout } \varphi \in C(\overline{\Omega}\times\mathbb{S}^{d-1}),$$

$$g^\varepsilon(x,e,m) = |e|^{\alpha_3} g\left(x, \frac{e}{|e|}, \frac{m}{|e|^{\alpha_4}}\right), \quad \textit{pour tout } \varepsilon > 0, (x,e) \in E^\varepsilon, m \geq 0,$$

et

$$\xi^\varepsilon(x,e) = \frac{t^\varepsilon(x,e)}{|e|^{\alpha_3}} \quad \textit{pour tout } (x,e) \in E^\varepsilon.$$

avec $\alpha_1, \ldots, \alpha_4$ *des réels qui vérifient la relation* $\alpha_1 + \alpha_4 = \alpha_2 - 1$*. Nous obtenons ainsi* (1.39) *où* I_0^ε *et* I_1^ε *sont donnés par :*

$$I_0^\varepsilon(\xi^\varepsilon) := \sum_{(x,e)\in E^\varepsilon} |e|^{\alpha_3+\alpha_4} H\left(x, \frac{e}{|e|}, \xi^\varepsilon(x,e)\right)$$

et

$$I_1^\varepsilon(\xi^\varepsilon) := \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y) \left(\min_{\sigma\in C_{x,y}^\varepsilon} \sum_{(z,e)\subset\sigma} |e|^{\alpha_3}\xi^\varepsilon(z,e)\right).$$

Alors au moins formellement, la limite continue de $J^\varepsilon(\xi^\varepsilon)/\varepsilon^{\alpha_3-\alpha_1-1}$ *est* $J(\xi)$ *donné par* (1.41)*. Dans les chapitres 2 et 3, nous avons choisi* $\alpha_1 = d/2-1, \alpha_2 = d$ *et* $\alpha_3 = \alpha_4 = d/2$*. Dans le chapitre 4, nous avons préféré* $\alpha_1 = 0, \alpha_2 = d, \alpha_3 = 1$ *et* $\alpha_4 = d-1$*. La relation* $\alpha_1+\alpha_4 = \alpha_2 - 1$ *est nécessaire pour que l'exposant dans* I_0^ε *soit cohérent avec celui de* I_1^ε*. La dimension d apparaît dans nos choix des exposants, en particulier, celui de* α_2 *(qui a toujours été choisi égal à d). Ainsi, l'exposant de* $I_0^\varepsilon/\varepsilon^{\alpha_3-\alpha_1-1}$ *est (formellement) égal à* α_2 $(= d)$*, ce qui nous permet de pouvoir passer à la limite quand* ε *tend vers* 0^+*.*

2.2 Conditions d'optimalité et variante à long terme

Ce chapitre reprend les deux dernières sections de [67]. Dans un premier temps, nous continuons de travailler dans le cas de court terme ($\Gamma = \{\gamma\}$) et nous gardons les mêmes hypothèses que dans le chapitre précédent. Nous voulons trouver des conditions d'optimalité pour le problème limite (1.41) à travers une formulation duale qui peut être vue comme un équilibre de Wardrop continu et qui fait donc appel à des mesures de probabilité sur l'ensemble des chemins. Cette question a été résolue dans le cas cartésien, toujours dans [12]. C'est une variante anisotropique du problème étudié dans [39]. Les auteurs ont considéré l'ensemble $C := W^{1,\infty}([0,1], \overline{\Omega})$, vu comme sous-ensemble de $C([0,1],\mathbb{R}^2)$ équipé de la topologie uniforme. L'ensemble des mesures de probabilité sur les chemins, compatibles avec le plan de transport γ, est :

$$\mathcal{Q}(\gamma) := \{Q \in \mathcal{M}_1^+(C) : (e_0, e_1)_\# Q = \gamma\}. \tag{1.42}$$

Pour $Q \in \mathcal{Q}(\gamma)$, l'analogue $m^Q = (m_1^Q, \dots, m_4^Q)$ de l'intensité de trafic i_Q (définition (1.8)) est défini par

$$\int_{\overline{\Omega}} \varphi \, dm_i^Q = \int_C \left(\int_0^1 \varphi(\sigma(t))(\dot\sigma(t) \cdot v_i)_+ dt \right) dQ(\sigma), \forall i = 1, \dots, 4 \text{ et } \varphi \in C(\overline{\Omega}, \mathbb{R}_+), \tag{1.43}$$

où ici, $(v_1, \dots, v_4) = ((1,0),(0,1),(-1,0),(0,-1))$, pour rappel. Compte-tenu de l'hypothèse (1.32), nous nous intéressons à ce sous-ensemble

$$\mathcal{Q}^q(\gamma) := \{Q \in \mathcal{Q}(\gamma) : m^Q \in L^q(\Omega, \mathbb{R}^4)\},$$

où q est l'exposant conjugué de p. L'équivalent du problème (1.20) est ainsi

$$\sup_{Q \in \mathcal{Q}^q(\gamma)} - \sum_{i=1}^4 \int_\Omega G_i(x, m_i^Q(x)) \, dx. \tag{1.44}$$

On a alors ce résultat, qui donne des conditions d'optimalité pour les problèmes (1.35) et (1.44) et qui fait le lien avec un équilibre de Wardrop continu :

Théorème 2.2. *Nous avons :*

1. *Le problème* (1.44) *admet des solutions.*
2. $\overline{Q} \in \mathcal{Q}^q(\gamma)$ *est une solution de* (1.44) *si et seulement si*

$$\sum_{i=1}^4 \int_C \left(\int_0^1 \xi_{\overline{Q}}(\sigma(t))(\dot\sigma(t) \cdot v_i)_+ dt \right) d\overline{Q}(\sigma) = \int_C \overline{c}_{\xi_{\overline{Q}}}(\sigma(0), \sigma(1)) \, d\overline{Q}(\sigma)$$

où $\xi_{\overline{Q}} := \left(g_1\left(\cdot, m_1^{\overline{Q}}(\cdot)\right), \dots, g_4\left(\cdot, m_4^{\overline{Q}}(\cdot)\right)\right)$.

3. *Nous avons l'égalité :* $\inf(1.35) = \sup(1.44)$. *De plus, si* $\overline{Q}$ *est une solution de* (1.44) *alors* $\xi_{\overline{Q}}$ *est une solution de* (1.35).

Dans le modèle général, nous devons considérer non seulement les chemins σ mais aussi les décompositions coniques de $\dot\sigma(t)$ dans la famille des directions $\{v_k(\sigma(t))\}$:

$$\mathcal{L} := \{(\sigma, \rho) : \sigma \in W^{1,\infty}([0,1], \overline{\Omega}), \rho \in \mathcal{P}_\sigma \cap L^\infty([0,1])^N\},$$

où

$$\mathcal{P}_\sigma := \left\{ \rho : t \in [0,1] \mapsto \rho(t) \in \mathbb{R}_+^N : \dot\sigma(t) = \sum_{k=1}^N v_k(\sigma(t)) \, \rho_k(t) \text{ p.p. } t \right\}.$$

Nous appelons ces (σ, ρ) des courbes généralisées. Les $\rho_k(t)$ sont les poids de $\dot\sigma(t)$ dans $\{v_k(\sigma(t))\}_k$. Cette construction nous permet notamment de traiter de manière différente une ligne droite et des oscillations autour de cette même ligne. En effet, dans le premier cas, nous avons une seule direction alors que dans le second cas, nous avons plusieurs directions. Par conséquent, nous utilisons des mesures de probabilité sur l'ensemble des courbes généralisées, qui sont cohérentes avec le plan de transport γ :

$$\mathcal{Q}(\gamma) := \{Q \in \mathcal{M}_+^1(\mathcal{L}) : (e_0, e_1)_\# Q = \gamma\},$$

Alors pour $Q \in \mathcal{Q}(\gamma)$, la mesure positive m^Q sur $\overline{\Omega} \times \mathbb{S}^{d-1}$ se définit comme suit :

$$\int_{\overline{\Omega} \times \mathbb{S}^{d-1}} \xi dm^Q = \sum_{k=1}^N \int_{\mathcal{L}} \left(\int_0^1 \xi(\sigma(t), v_k(\sigma(t)))\rho_k(t) \, dt \right) dQ(\sigma, \rho), \forall \xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$$

et pour la même raison que dans le cas cartésien, nous nous concentrons sur l'ensemble (que nous supposons non vide)

$$\mathcal{Q}^q(\gamma) := \{Q \in \mathcal{Q}(\gamma) \ : \ m^Q \in L^q(\Omega \times \mathbb{S}^{d-1}, \theta)\}. \tag{1.45}$$

Alors la version générale du problème (1.44) est

$$\sup_{Q \in \mathcal{Q}^q(\gamma)} - \int_{\Omega \times \mathbb{S}^{d-1}} G\left(x, v, m^Q(x,v)\right) \, \theta(dx, dv). \tag{1.46}$$

Le résultat principal de cette section est ainsi le théorème suivant

Théorème 2.3. *Sous certaines hypothèses, nous avons :*

1. (1.46) *admet des solutions.*
2. $\overline{Q} \in \mathcal{Q}^q(\gamma)$ *est une solution de* (1.46) *si et seulement si*

$$\sum_{k=1}^N \int_{\mathcal{L}} \left(\int_0^1 \xi(\sigma(t), v_k(\sigma(t)))\rho_k(t) \, dt \right) d\overline{Q}(\sigma, \rho) = \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0), \sigma(1)) \, d\overline{Q}(\sigma, \rho)$$

où $\xi_{\overline{Q}}(x, v) = g\left(x, v, m^{\overline{Q}}(x,v)\right)$.

3. *Nous avons l'égalité :* $\inf (1.41) = \sup (1.46)$. *De plus, si* $\overline{Q}$ *est une solution de* (1.46) *alors* $\xi_{\overline{Q}}$ *est une solution de* (1.41).

Cependant, la preuve du théorème et en particulier de l'existence d'une solution optimale s'avère beaucoup plus délicate. En effet, nous devons composer avec des courbes généralisées, qui ne sont pas dans un espace polonais (c'est-à-dire, un espace métrisable, séparable et complet), quelle que soit la topologie utilisée. Ce fait nous empêche ainsi d'appliquer directement le théorème de Prokhorov, qui aurait permis de prendre une suite maximisante $\{Q_n\}_{n\geq 0}$ pour (1.46). L'idée est de contourner cette difficulté en considérant une version relaxée du problème (1.46). Plus précisément, partant du problème (1.46), nous étendons la classe des objets sur laquelle le supremum est pris, en plongeant l'espace $\mathcal{L}$ dans l'espace

$$\mathcal{S} = \left\{ (\sigma, \nu_t \otimes \lambda) : \sigma \in W^{1,\infty}([0,1], \overline{\Omega}), \nu_t \in \mathfrak{M}_\sigma^t \text{ p.p. } t \right\},$$

où pour $t \in [0,1]$ et $\sigma \in W^{1,\infty}([0,1], \overline{\Omega})$,

$$\mathfrak{M}_\sigma^t := \left\{ \nu_t \in \mathcal{M}_+^1(\mathbb{R}^d) : \text{supp } \nu_t \subset \bigcup_{k=1}^N \mathbb{R}_+ v_k(\sigma(t)) \text{ et } \dot{\sigma}(t) = \int_{\mathbb{R}^d} v \, d\nu_t(v) \right\}$$

et λ est la mesure de Lebesgue sur $[0,1]$. Les mesures d'Young $\nu_t \otimes \lambda$ sont l'équivalent des décompositions $\rho \in \mathcal{P}_\sigma$. La théorie des mesures d'Young a été introduite par Young dans [99–101]. Une référence est le livre de Pedregal [85]. Nous avons vu au début de ce chapitre que le problème de Kantorovich (1.3) est une version relaxée du problème de Monge (1.1). Pour une comparaison entre les idées de Kantorovich et celles d'Young, nous pouvons consulter les articles [10, 62].

L'ensemble $\mathcal{S}$ est vu comme sous-ensemble de $C = C([0,1], \mathbb{R}^d) \times \mathfrak{P}_1(\mathbb{R}^d \times [0,1])$ où pour un espace polonais (E, d), nous posons

$$\mathfrak{P}_1(E) := \left\{ \mu \in \mathcal{M}_+^1(E) : \int_E \mathrm{d}(x, x') \, d\mu(x) < +\infty \ \text{ pour } x' \in E \right\}.$$

Nous munissons $C([0,1],\mathbb{R}^d)$ de la topologie uniforme et $\mathfrak{P}_1(\mathbb{R}^d \times [0,1])$ de celle induite par la 1-distance de Wasserstein

$$W_1(\mu,\nu) := \min\left\{\int_{E^2} \mathrm{d}(x_1,x_2)\, d\pi(x_1,x_2) : \pi \in \Pi(\mu,\nu)\right\}$$

où $E = \mathbb{R}^d \times [0,1]$, d est la distance usuelle sur E et $(\mu,\nu) \in \mathfrak{P}_1(E)^2$. Nous équipons alors C de la topologie-produit et C est ainsi un espace polonais. Nous pouvons remarquer le lien entre W_1 et le problème de Kantorovich (1.3). Une bonne référence sur les distances de Wasserstein est Ambrosio-Gigli-Savaré [9].

Nous pouvons réécrire le problème (1.46) en un problème de minimisation sur des mesures de probabilité sur l'ensemble $\mathcal{S}$. En outre, une hypothèse faite au début de [67] et jouant un rôle crucial dans cette preuve est :

Hypothèse 2.3. *Il existe une constantc $C > 0$ telle que pour tout $(x,z,\xi) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}_+^N$, il existe $\bar{Z} \in \mathbb{R}_+^N$ tel que $|\bar{Z}| \le C$ et*

$$\bar{Z} \cdot \xi = \min\left\{Z \cdot \xi; Z = (z_1,\dots,z_N) \in \mathbb{R}_+^N \text{ et } \sum_{k=1}^N z_k v_k(x) = z\right\}.$$

Cette hypothèse signifie que pour tout $z \in \mathbb{R}^d$, il existe une décomposition conique minimale de z dans la famille des directions qui ne soit pas trop grande par rapport à z. Elle nous permet de réduire nos problèmes de minimisation en ne considérant que des mesures de probabilité sur des ensembles plus petits que $\mathcal{L}$ et $\mathcal{S}$, sur lesquels nous avons une contrainte supplémentaire de contrôle. Cela nous donne de la tension sur les mesures, ce qui nous permet d'appliquer le théorème de Prokhorov.

Dans la dernière section, nous travaillons avec la variante de long-terme. Dans les problèmes discrets, au lieu de prendre des plans de transport $\{\gamma^\varepsilon\}_{\varepsilon>0}$, nous fixons seulement les marginales $\{f_-^\varepsilon\}_{\varepsilon>0}$ et $\{f_+^\varepsilon\}_{\varepsilon>0}$ qui convergent $\star$-faiblement vers des mesures de probabilité f_- et f_+ sur $\overline{\Omega}$. Les problèmes de minimisation discrets (1.39) et continu (1.41) se reformulent ainsi :

$$\inf_{\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}} F^\varepsilon(\xi^\varepsilon) := I_0^\varepsilon(\xi^\varepsilon) - F_1^\varepsilon(\xi^\varepsilon) \tag{1.47}$$

où $I_0^\varepsilon(\xi^\varepsilon)$ est défini par (1.40) et

$$F_1^\varepsilon(\xi^\varepsilon) := \inf_{\gamma^\varepsilon \in \Pi(f_-^\varepsilon, f_+^\varepsilon)} \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y) \left(\min_{\sigma \in C_{x,y}^\varepsilon} \sum_{(z,e)\subset\sigma} |e|^{d/2} \xi^\varepsilon(z,e)\right).$$

et

$$F(\xi) := I_0(\xi) - F_1(\xi), \text{ où } F_1(\xi) := \inf_{\gamma\in\Pi(f_-,f_+)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi d\gamma,\ \forall \xi \in L_+^p(\theta). \tag{1.48}$$

Sous les mêmes hypothèses que dans le cas de court-terme (excepté celle concernant le plan de transport, remplacée par une autre sur les martingales), nous avons le même résultat de Γ-convergence : la famille de fonctionnelles (1.47) Γ-converge (pour la topologie faible de L^p) vers la fonctionnelle (1.48).

De plus, au lieu de prendre $\mathcal{Q}^q(\gamma)$ (définition (1.45)), nous travaillons maintenant avec

$$\mathcal{Q}^q(f_-, f_+) := \{Q \in \mathcal{M}^1_+(\mathcal{L}) : e_{0\#}Q = f_-, e_{1\#}Q = f_+, m^Q \in L^q(\theta)\}$$

et le problème (1.46) devient

$$\sup_{Q\in\mathcal{Q}^q(f_-,f_+)} - \int_{\Omega\times\mathbb{S}^{d-1}} G(x, v, m^Q(x, v))\theta(dx, dv).$$

En faisant les modifications nécessaires (voir chapitre 3), le Théorème 2.3 reste alors encore vrai dans le cas de long terme.

2.3 Équilibre de Wardrop : variante à long terme, EDPs dégénérées et anisotropiques et approximations numériques

Ce chapitre est issu de l'article [66]. C'est la suite de [67] et nous reprenons pratiquement les mêmes notations et définitions. À la fin de [67], nous avions le problème de minimisation (dans le cadre continu de long terme)

$$\inf_{Q\in\mathcal{Q}^q(f_-,f_+)} \int_{\Omega\times\mathbb{S}^{d-1}} G(x, v, m^Q(x, v))\theta(dx, dv). \tag{1.49}$$

Considérons le problème à la Beckmann suivant :

$$\inf_{\sigma\in L^q(\Omega,\mathbb{R}^d)} \left\{\int_\Omega \mathcal{G}(x, \sigma(x))\, dx : \; -\text{div}\ \sigma = f\right\}. \tag{1.50}$$

où $f = f_+ - f_-$ et $\mathcal{G} : \Omega\times\mathbb{R}^d \to \mathbb{R}_+$ est la fonction, convexe par rapport à la seconde variable, définie par

$$\mathcal{G}(x, \sigma) := \inf_{\varrho\in\mathcal{P}^\sigma_x} \sum_{k=1}^N c_k(x)G(x, v_k(x), \varrho_k)$$

avec

$$\mathcal{P}^\sigma_x := \left\{\varrho \in \mathbb{R}^N_+;\ \sigma = \sum_{k=1}^N v_k(x)\varrho_k\right\} \text{ pour tous } x \in \Omega \text{ et } \sigma \in \mathbb{R}^d.$$

Nous avons le théorème suivant :

Théorème 2.4. *Nous avons l'égalité* inf (1.49) = inf (1.50).

Il généralise un théorème de Brasco-Carlier-Santambrogio [33] dans le cas isotrope et de Brasco-Carlier [31] dans le cas cartésien. Pour prouver la première inégalité inf (1.49) $\geq$ inf (1.50), à partir de $Q \in \mathcal{Q}^q(f_-, f_+)$, nous construisons la mesure vectorielle $\sigma^Q \in L^q(\Omega, \mathbb{R}^d)$ de la même manière que dans la définition (1.9) et nous obtenons notamment que $(m^Q(\cdot, v_1(\cdot)), \ldots, m^Q(\cdot, v_N(\cdot))) \in \mathcal{P}^{\sigma^Q}$ et l'inégalité désirée s'ensuit. Pour l'inégalité inverse, nous utilisons la méthode de flot de Moser et un argument de régularisation classique.

Comme déjà mentionné auparavant, par des arguments standards de dualité convexe (le théorème de Fenchel-Rockafellar, [54]), le problème dual de (1.50) est

(1.23) et nous avons min (1.50) = min (1.23). De plus, nous pouvons caractériser l'élément minimisant σ par (1.24). Ainsi, résoudre (1.50) revient à résoudre l'équation d'Euler-Lagrange de (1.23) et à utiliser ensuite les conditions d'optimalité primal-dual. Cependant, ici, dans nos modèles de congestion, même en l'absence totale de trafic, nous ne pouvons pas aller à une vitesse infinie donc la dérivée de la fonction $G(x, v, \cdot)$, qui est $g(x, v, 0)$, est strictement positive en zéro. Cela rend l'équation de Euler-Lagrange très dégénérée. Un exemple typique est $g(x, v_k(x), m) = g_k(x, m) = m^{q-1} + \delta_k$ avec $\delta_k > 0$. Dans le cas cartésien, cette équation devient

$$-\sum_{k=1}^{2} \partial_k \left((|\partial_k u| - \delta_k)_+^{p-1} \frac{\partial_k u}{|\partial_k u|} \right) = f,$$

Dans le cas général, elle est encore plus compliquée :

$$-\sum_{l=1}^{d} \partial_l \left[\sum_{k=1}^{N} (\nabla u \cdot v_k(x) - \delta_k c_k(x))_+^{p-1} v_k^l(x) \right] = f.$$

où $v_k(x) = (v_k^1(x), \ldots, v_k^d(x))$ pour $k = 1, \ldots, N$ et $x \in \overline{\Omega}$. Cette EDP est très dégénérée. En effet, tout u, dont, pour tout x, le gradient $\nabla u(x)$ est à valeurs dans un polyèdre (non trivial et qui dépend de x), est une solution de l'équation précédente avec $f = 0$. Par conséquent, nous ne pouvons pas espérer récupérer des estimations sur les dérivées secondes de u ou même des estimations d'oscillations sur ∇u à partir de cette EDP. Même en considérant le cas cartésien avec les δ_k tous nuls, l'équation (qui est alors l'équation du pseudo p-laplacien, voir par exemple [18, 77, 96, 97]) demeure délicate à étudier.

Dans le cas particulier où les directions v_k et les coefficients de volume c_k sont constants, le problème (1.50) s'écrit plus simplement

$$\inf_{\sigma \in L^q(\Omega)} \left\{ \int_\Omega \inf_{\varrho \in \mathcal{P}_x^\sigma} \sum_{k=1}^{N} c_k \left(\frac{1}{q} \varrho_k^q + \delta_k \varrho_k \right) : -\text{div } \sigma = f \right\}, \tag{1.51}$$

et nous avons le résultat suivant de régularité sur σ :

Corollaire 2.2. *La solution σ de* (4.25) *est dans l'espace de Sobolev $W^{1,r}_{loc}(\Omega)$, où*

$$r = \begin{cases} 2 & si\ p = 2, \\ toute\ valeur\ < 2, & si\ p > 2\ et\ d = 2, \\ \dfrac{dp}{dp - (d+p) + 2}, & si\ p > 2\ et\ d > 2. \end{cases}$$

Le même résultat a été prouvé pour le cas cartésien dans [31]. Les preuves sont très semblables et sont basées sur la méthode des translations de Nirenberg (voir par exemple Brézis [37], Gilbarg-Trudinger [63], Lindqvist [76]) et des inégalités pour le p-laplacien [76].

Dans la dernière section, nous approchons numériquement par la méthode des éléments finis les solutions du problème de minimisation (1.23). Pour cela, nous utilisons l'algorithme ALG2 comme dans Benamou-Carlier [21], qui est un cas particulier de la méthode de *splitting* de Douglas-Rachford pour la somme de deux opérateurs non-linéaires (voir Lions-Mercier [78] ou encore Papadakis-Peyré-Oudet [84]). Cet algorithme a été introduit par Fortin-Glowinski [59]. Malgré la relative lenteur de sa convergence, cette méthode itérative fonctionne bien numériquement. Il existe à présent une énorme littérature concernant les applications de ALG2. Pour le transport, citons notamment Benamou-Brenier [20], Benamou [19], Buttazzo-Jimenez-Oudet [38], Glowinski [64], Glowinski-Marocco [65], Huilgol-You [69]. Le problème discrétisé par éléments finis de (1.23) est :

$$\inf_{u\in\mathbb{R}^n} J(u) := \mathbf{F}(u) + \mathbf{G}^*(\Lambda u) \tag{1.52}$$

où $\mathbf{F} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \mathbf{G} : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ sont des fonctions convexes, semi-continues inférieurement et propres, Λ est une $m \times n$ matrice réelle, représentant l'analogue discrète de ∇. Par ailleurs, nous pouvons réécrire (1.23) :

$$\inf_{u,q} \left\{ \int_\Omega \mathcal{G}^*(x, q(x))\, dx - \int_\Omega u(x) f(x)\, dx \right\}. \tag{1.53}$$

par rapport à la contrainte $\nabla u = q$. Alors nous pouvons reformuler les problèmes (1.53)-(1.50) en un problème de point-selle :

$$\inf_{u,q} \sup_\sigma L_r(u, q, \sigma)$$

pour $r > 0$, où le Lagrangien augmenté L_r est défini par

$$L_r(u, q, \sigma) := \int_\Omega \mathcal{G}^*(x, q(x))\, dx - \langle u, f\rangle + \langle \sigma, \nabla u - q\rangle + \frac{r}{2}|\nabla u - q|^2.$$

La version discrète de ce Lagrangien augmenté est :

$$L_r(u, q, \sigma) := \mathbf{F}(u) + \mathbf{G}^*(q) + \sigma \cdot (\Lambda u - q) + \frac{r}{2}|\Lambda u - q|^2,\ \forall (u, q, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m. \tag{1.54}$$

L'algorithme de Lagrangien augmenté ALG2 consiste à construire $(u^k, q^k, \sigma^k) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ à partir d'une donnée initiale (u^0, q^0, σ^0) comme suit :

1. Problème de minimisation par rapport à u :

   $$u^{k+1} := \mathrm{argmin}_{u\in\mathbb{R}^n} \left\{ \mathbf{F}(u) + \sigma^k \cdot \Lambda u + \frac{r}{2}|\Lambda u - q^k|^2) \right\}$$

   C'est équivalent à résoudre une équation de Laplace :

   $$-r(\Delta u^{k+1} - \mathrm{div}(q^k)) = f + \mathrm{div}(\sigma^k) \text{ dans } \Omega$$

   avec la condition de bord de Neumann

   $$r\frac{\partial u^{k+1}}{\partial \nu} = rq^k \cdot \nu - \sigma^k \cdot \nu \text{ sur } \partial\Omega.$$

2. Problème de minimisation par rapport à q :

   $$q^{k+1} := \mathrm{argmin}_{q\in\mathbb{R}^d} \left\{ \mathbf{G}^*(q) - \sigma^k \cdot q + \frac{r}{2}|\Lambda u^{k+1} - q|^2) \right\}$$

3. Utilisation de la formule de remontée de σ

   $$\sigma^{k+1} := \sigma^k + r(\Lambda u^{k+1} - q^{k+1}).$$

La convergence d'une telle suite (u^k, q^k, σ^k) vers un point-selle du Lagrangien est assurée par le théorème de Eckstein-Bertsekas [53].

Le logiciel utilisé pour l'implémentation de cet algorithme est FreeFem ++ [68]. Je l'ai testé dans deux cas particuliers dans $\mathbb{R}^2$: le cartésien et le hexagonal. La seule chose qui change entre ces deux cas est l'étape 2, les deux autres étapes étant identiques. Dans le premier cas, cela revient à trouver la solution $q = (q_1, q_2)$ du problème ponctuel

$$\inf_{q_i} \frac{1}{p}(|q_i| - c(x))_+^p + \frac{r}{2}|q_i - \tilde{q}_i^k|^2 \quad \text{pour } i = 1, 2$$

où $\tilde{q}^k = \Lambda u^{k+1} + \frac{\sigma^k}{r}$. Une simple dichotomie suffit. En revanche, c'est plus délicat dans le second cas, comme nous ne pouvons pas séparer les variables dans $\mathbf{G}^*(q)$. Nous cherchons l'élément minimisant avec la méthode de Newton, en utilisant l'inverse de la matrice hessienne (qui est définie positive). Des tests ont été effectués avec plusieurs données différentes : f, p, présence d'un obstacle, etc. La convergence de cette discrétisation a été vérifiée dans tous les tests grâce à trois critères de convergence.

2.4 Une solution numérique au problème de Monge avec comme coût une distance de Finsler

Dans ce chapitre écrit en collaboration avec Benamou et Carlier [22], nous nous intéressons au problème de Monge

$$\inf \left\{ J(\pi) := \int_{X \times Y} c(x, y)\, d\pi(x, y) : \pi \in \Pi(\mu, \nu) \right\}, \tag{1.55}$$

où Ω est un domaine borné de $\mathbb{R}^d$, f^- et f^+ sont des mesures de probabilité sur $\overline{\Omega}$ et d_L est une distance de Finsler. Plus précisément, d_L est donnée par

$$d_L(x, y) := \inf \left\{ \int_0^1 L(\gamma(t), \dot{\gamma}(t))\ \mathrm{d}t \ : \ \gamma \in W^{1,1}([0, 1], \overline{\Omega}),\ \gamma(0) = x,\ \gamma(1) = y \right\} \tag{1.56}$$

où le Lagrangien $L : \overline{\Omega} \times \mathbb{R}^d \to \mathbb{R}_+$ est une fonction continue de type Finsler, c'est-à-dire, pour tout $x \in \overline{\Omega}$, $v \mapsto L(x, v)$ est une norme et il existe une constante $C > 0$ telle que la condition suivante de non-dégénérescence est satisfaite :

$$\frac{|v|}{C} \leq L(x, v) \leq C|v|,\ \forall (x, v) \in \overline{\Omega} \times \mathbb{R}^d.$$

Une référence générale sur les distances de Finsler est le livre de Bao-Chern-Shen [14]. Nous pouvons également consulter Braides-Buttazzo-Fragalà [30] pour une approximation riemannienne des métriques de Finsler. Nous pouvons reformuler le problème (1.56) en un problème à la Kantorovich :

$$\sup \left\{ \langle u, f \rangle := \int_\Omega u(x) \mathrm{d}(f^+ - f^-)(x) \ : \ u \text{ est 1-Lipschitzienne pour } d_L \right\}. \tag{1.57}$$

et le problème dual de flot à la Beckmann est alors :

$$\inf_{\sigma \in L^1(\Omega, \mathbb{R}^d)} \left\{ \int_\Omega L(x, \sigma(x)) \mathrm{d}x \ : \ -\operatorname{div}(\sigma) = f \right\}. \tag{1.58}$$

Par dualité, nous montrons que

$$\min(1.55) = \max(1.57) = \min(1.58).$$

Comme dans le chapitre précédent, nous réécrivons (1.57) en

$$\inf_{u,q}\Big\{ -\langle u, f\rangle + G(q) \ : \ q = \nabla u \text{ p.p.}\Big\}$$

où

$$G(q) := \int_\Omega \mathcal{G}(x, q(x))\mathrm{d}x$$

et

$$\mathcal{G}(x,q) := \begin{cases} 0, \ \text{si } L^*(x,q) \le 1 \\ +\infty, \text{sinon.}\end{cases}$$

Ici, $L^*(x,\cdot)$ est la norme duale de $L(x,\cdot)$:

$$L^*(x,p) := \sup\{p\cdot v \ : \ L(x,v) \le 1\}.$$

Comme précédemment, nous considérons le Lagrangien augmenté (pour $r > 0$)

$$\begin{aligned} L_r(u,q,\sigma) := &-\langle u, f\rangle + \int_\Omega \mathcal{G}(x,q(x))\mathrm{d}x \\ &+ \int_\Omega \sigma(x)\cdot(\nabla u(x) - q(x))\mathrm{d}x + \frac{r}{2}\int_\Omega |\nabla u - q|^2. \end{aligned}$$

De la même manière, nous faisons des approximations discrètes de ces problèmes en utilisant les éléments finis pour pouvoir appliquer l'algorithme ALG2 et faire des simulations numériques. À notre connaissance, il n'existait pas de littérature sur l'utilisation des méthodes de Lagrangien augmenté pour une métrique de Finsler générale. Nous justifions rigoureusement la pertinence d'une telle approche. Nous testons l'algorithme sur deux exemples particuliers : le cas riemannien

$$L(x,v) := (A(x)v\cdot v)^{\frac{1}{2}},$$

avec $A(x)$ matrice symétrique définie positive pour tout x et le cas où $L(x,\cdot)$ est défini par un nombre fini de directions

$$L(x,v) := \inf\left\{\sum_{j=1}^{2k}\xi_j(x)\alpha_j \ : \ \sum_{j=1}^{2k}\alpha_j v_j(x) = v\right\},$$

où $\{v_j(x)\}$ sont des vecteurs unité et $\{\xi(x)\}_j$ sont des réels strictement positifs tels que $v_{j+k}(x) = -v_j(x)$ et $\xi_{j+k} = \xi_j$ pour $j \in \{1, \cdots, k\}$.

En résumé, les chapitres 2 et 3 reprennent Hatchi [67]. Dans le chapitre 2, nous regardons comment se passe la transition d'une suite de réseaux congestionnés discrets de plus en plus denses à un modèle continu dans le cas de court terme. Dans le chapitre 3, nous donnons des conditions d'optimalité dans le modèle continu. Le chapitre 4 est issu de Hatchi [66] et poursuit l'étude du modèle continu dans le cas de long terme. Nous arrivons ainsi à une EDP anisotropique, elliptique et dégénérée et nous la résolvons grâce à l'algorithme ALG2. Le chapitre 5, écrit par Benamou-Carlier-Hatchi [22], porte sur les problèmes de Monge avec comme coût une distance de Finsler et nous utilisons à nouveau l'algorithme ALG2 pour avoir une solution numérique. Le dernier chapitre présente des perspectives possibles pour la poursuite des travaux exposés dans cette thèse.

Chapter 2

Wardrop equilibria : rigorous derivation of continuous limits from general networks models

This chapter is the first part of the paper [67].

Abstract : The concept of Wardrop equilibrium plays an important role in congested traffic problems since its introduction in the early 50's. As shown in [12], when we work in two-dimensional cartesian and increasingly dense networks, passing to the limit by Γ-convergence, we obtain continuous minimization problems posed on measures on curves. Here we study the case of general networks in $\mathbb{R}^d$ which become very dense. We use the notion of generalized curves and extend the results of the cartesian model.

Keywords: traffic congestion, Wardrop equilibrium, Γ-convergence, generalized curves.

1 Introduction

Modeling congested traffic is a field of research that has developed especially since the early 50's and the introduction of Wardrop equilibrium [98]. Its popularity is due to many applications to road traffic and more recently to communication networks. In our finite networks model, we represent the congestion effects by the fact that the traveling time of each arc is a nondecreasing function of the flow on this arc. The concept of Wardrop equilibrium simply says that all used roads between two given points have the same cost and this cost is minimal. So we assume a rational behavior by users. A Wardrop equilibrium is a flow configuration that satisfies mass conservation conditions and positivity constraints. A few years after Wardrop defined his equilibrium notion, Beckmann, McGuire and Winsten [17] observed that Wardrop equilibrium can be formulated in terms of solutions of a convex optimization problem. However this variational characterization uses the whole path flow configuration. It becomes very costly when working in increasingly dense networks. We may often prefer to study the dual problem which is less untractable. But finding an optimal solution remains a hard problem because of the presence of a nonsmooth and nonlocal term. As we study a sequence of discrete networks increasingly dense, it is natural to ask under what conditions we can pass to a continuous limit which would simplify the problem.

The purpose of this paper is to rigorously justify passing to the limit thanks to the theory of Γ-convergence and then to find a continuous analogue of Wardrop equilibrium. We will strongly rely on two articles [39] and [12]. The first establishes some first results on traffic congestion. The second studies the case of a two-dimensional cartesian grid with small arc length ε. Here we will consider general networks in $\mathbb{R}^d$ with small arc length of order ε. It is a substantial improvement. We will show the Γ-convergence of the functionals in the dual problem as ε goes to 0. We will obtain an optimization problem over metrics variables. The proof of the Γ-convergence is constructed in the same manner as in [12]. But two major difficulties here appear. Indeed in the case of the grid in $\mathbb{R}^2$, there are only four possible constant directions (that are $((1,0),(0,1),(-1,0),(0,-1))$) so that for all speed $z \in \mathbb{R}^2$, there exists an unique decomposition of z in the family of these directions, with positive coefficients. In the general case, directions are not necessarily constant and we have no uniqueness of the decomposition. To understand how to overcome these obstacles, we can first look at the case of regular hexagonal networks. There are six constant directions (that are $\exp(i(\pi/6+k\pi/3)), k=0,\dots,5$) but we lose uniqueness. Then, we can study the case of a two-dimensional general network in which directions can vary and arcs lengths are not constant. The generalization from $\mathbb{R}^2$ to $\mathbb{R}^d$ (where d is any integer ≥ 2) is simpler. Of course, it is necessary to make some structural assumptions on the networks to have the Γ-convergence. These hypotheses are satisfied for instance in the cases of the isotropic model in [39] and the cartesian one in [12].

The limit problem (in the general case) is the dual of a continuous problem posed on a set of probability measures over generalized curves of the form (σ, ρ) where σ is a path and ρ is a positive decomposition of $\dot{\sigma}$ in the family of the directions. This takes the anisotropy of the network into account. We will then remark that we can define a continuous Wardrop equilibrium through the optimality conditions for the continuous model. To establish that the limit problem has solutions, we work with a relaxation of this problem through the Young's measures and we extend results in [39]. Indeed we cannot directly apply Prokhorov's theorem to the set of generalized curves since weak-L^1 is not contained in a Polish space. First we are interested in the short-term problem, that is, we have a transport plan that gives the amount of mass sent from each source to each destination. We may then generalize these results to the long-term variant in which only the marginals (that are the distributions of supply and demand) are known. This case is interesting since as developed in [31, 33, 66], it amounts to solve a degenerate elliptic PDE. But it will not be developed here.

The plan of the paper is as follows: Section 2 is devoted to a preliminary description of the discrete model with its notations, definition of Wardrop equilibrium and its variational characterization. In Section 3, we explain the assumptions made and we identify the limit functional. We then state the Γ-convergence result. The proof is given in Section 4. Then, in Section 5, we formulate the optimality conditions for the limit problem that lead to a continuous Wardrop equilibrium. Finally, in Section 6, we adapt the previous results to the long-term problem.

2 The discrete model

2.1 Notations and definition of Wardrop equilibria

Let $d \in \mathbb{N}, d \geq 2$ and Ω a bounded domain of $\mathbb{R}^d$ with a smooth boundary and $\varepsilon > 0$. We consider a sequence of discrete networks $\Omega_\varepsilon = (N^\varepsilon, E^\varepsilon)$ whose characteristic length is ε, where N^ε is the set of nodes in Ω_ε and E^ε the (finite) set of pairs (x, e) with $x \in N^\varepsilon$ and $e \in \mathbb{R}^d$ such that the segment $[x, x + e]$ is included in Ω and $x + e$ still belongs to N^ε. We will simply identify arcs to pairs (x, e). We impose $|E^\varepsilon| = \max\{|e|, \text{ there exists } x \text{ such that } (x, e) \in E^\varepsilon\} = \varepsilon$. We may assume that two arcs can not cross. The orientation is important since the arcs (x, e) and $(x + e, -e)$ really represent two distinct arcs. Moreover if $(x, e) \in E^\varepsilon$ then also is $(x + e, -e)$. Now let us give some definitions and notations.

Traveling times and congestion: We denote the mass commuting on arc (x, e) by $m^\varepsilon(x, e)$ and the traveling time of arc (x, e) by $t^\varepsilon(x, e)$. We represent congestion by the following relation between traveling time and mass for every arc (x, e):

$$t^\varepsilon(x, e) = g^\varepsilon(x, e, m^\varepsilon(x, e)) \tag{2.1}$$

where for every ε, g^ε is a given positive function that depends on the arc itself but also on the mass $m^\varepsilon(x, e)$ that commutes on the arc (x, e) in a nondecreasing way: this is congestion. We will denote the set of all arc-masses $m^\varepsilon(x, e)$ by $\mathbf{m}^\varepsilon$. Orientation of networks here is essential: considering two neighboring nodes x and x' with $(x, x' - x)$ and $(x', x - x') \in E^\varepsilon$, the time to go from x to x' only depends on the mass $m^\varepsilon(x, x' - x)$ that uses the arc $(x, x' - x)$ whereas the time to go from x' to x only depends on the mass $m^\varepsilon(x', x - x')$.

Transport plan: A transport plan is a given function $\gamma^\varepsilon : N^\varepsilon \times N^\varepsilon \mapsto \mathbb{R}_+$. That is a collection of nonnegative masses where for each pair $(x, y) \in N^\varepsilon \times N^\varepsilon$ (viewed as a source/destination pair), $\gamma^\varepsilon(x, y)$ is the mass that has to be sent from the source x to the target y.

Paths: A path is a finite collection of successive nodes. We therefore represent a path σ by writing $\sigma = \left(x_0, \dots, x_{N(\sigma)}\right)$ with $\sigma(k) = x_k \in N^\varepsilon$ and $(\sigma(k), \sigma(k+1) - \sigma(k)) \in E^\varepsilon$ for $k = 0, \dots, N(\sigma) - 1$. The node $\sigma(0)$ is the origin of σ and $\sigma(N(\sigma))$ is the terminal point of σ. The length of σ is

$$\sum_{k=0}^{N(\sigma)-1} |x_{k+1} - x_k|.$$

We say that $(x, e) \subset \sigma$ if there exists $k \in \{1, \dots, N(\sigma) - 1\}$ such that $\sigma(k) = x$ and $e = \sigma(k+1) - \sigma(k)$. Since the time to travel on each arc is positive, we can impose σ has no loop. We will denote the (finite) set of loop-free paths by C^ε, that may be partitioned as

$$C^\varepsilon = \bigcup_{(x,y) \in N^\varepsilon \times N^\varepsilon} C^\varepsilon_{x,y},$$

where $C^\varepsilon_{x,y}$ is the set of loop-free paths starting from the origin x and stopping at the terminal point y. The mass commuting on the path $\sigma \in C^\varepsilon$ will be denoted $w^\varepsilon(\sigma)$. The collection of all path-masses $w^\varepsilon(\sigma)$ will be denoted $\mathbf{w}^\varepsilon$. Given arc-masses $\mathbf{m}^\varepsilon$, the travel time of a path $\sigma \in C^\varepsilon$ is given by:

$$\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma) = \sum_{(x,e) \subset \sigma} g^\varepsilon(x, e, m^\varepsilon(x, e)).$$

Equilibria: In short, in this model, the data are the masses $\gamma^\varepsilon(x, y)$ and the congestion functions g^ε. The unknowns are the arc-masses $m^\varepsilon(x, e)$ and path-masses $w^\varepsilon(\sigma)$. We wish to define some equilibrium requirements on these unknowns. First, they should be nonnegative. Moreover, we have the following conditions that relate arc-masses, path-masses and the data γ^ε :

$$\gamma^\varepsilon(x, y) := \sum_{\sigma \in C^\varepsilon_{x,y}} w^\varepsilon(\sigma), \ \forall (x, y) \in N^\varepsilon \times N^\varepsilon \tag{2.2}$$

and

$$m^\varepsilon(x, e) = \sum_{\sigma \in C^\varepsilon : (x,e) \subset \sigma} w^\varepsilon(\sigma), \forall (x, e) \in E^\varepsilon. \tag{2.3}$$

Both express mass conservation. We finally require that only the shortest paths (taking into account the congestion created by arc and path-masses) should actually be used. This is the concept of Wardrop equilibrium that is defined precisely as follows:

Definition 2.1. *A Wardrop equilibrium is a configuration of nonnegative arc-masses $\mathbf{m}^\varepsilon : (x, e) \to (m^\varepsilon(x, e))$ and of nonnegative path-masses $\mathbf{w}^\varepsilon : \sigma \to w^\varepsilon(\sigma)$, satisfying the mass conservation conditions* (2.2) *and* (2.3) *and such that for every $(x, y) \in N^\varepsilon \times N^\varepsilon$ and every $\sigma \in C^\varepsilon_{x,y}$, if $w^\varepsilon(\sigma) > 0$ then*

$$\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma) \leq \tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma'), \forall \sigma' \in C^\varepsilon_{x,y}.$$

2.2 Variational characterizations of equilibria

Soon after the work of Wardrop, Beckmann, McGuire and Winsten [17] discovered that Wardrop equilibria can be obtained as minimizers of a convex optimization problem:

Theorem 2.1. *A flow configuration* $(\mathbf{w}^\varepsilon, \mathbf{m}^\varepsilon)$ *is a Wardrop equilibrium if and only if it minimizes*

$$\sum_{(x,e)\in E^\varepsilon} G^\varepsilon(x,e,m^\varepsilon(x,e)) \text{ where } G^\varepsilon(x,e,m) := \int_0^m g^\varepsilon(x,e,\alpha)d\alpha \tag{2.4}$$

subject to nonnegativity constraints and the mass conservation conditions (2.2)-(2.3).

Proof. Problem (2.4) in fact is a minimization problem only on $\mathbf{w}^\varepsilon$ since we can deduce $\mathbf{m}^\varepsilon$ from $\mathbf{w}^\varepsilon$ due to (2.3). Assume that $\mathbf{w}^\varepsilon$ (with associated arc-masses $\mathbf{m}^\varepsilon$) is optimal for (2.4) then for every admissible $v = (v(\sigma))_{\sigma\in C^\varepsilon}$ with associated arc-masses $n = (n(x,e))_{(x,e)\in E^\varepsilon}$ (through (2.3)), we have

$$\begin{aligned} 0 &\leq \sum_{(x,e)\in E^\varepsilon} G^{\varepsilon\prime}(x,e,m^\varepsilon(x,e))(n(x,e) - m^\varepsilon(x,e)) \\ &= \sum_{(x,e)\in E^\varepsilon} g^\varepsilon(x,e,m^\varepsilon(x,e)) \sum_{\sigma\in C^\varepsilon:(x,e)\in\sigma} (v(\sigma) - w^\varepsilon(\sigma)) \\ &= \sum_{\sigma\in C^\varepsilon} (v(\sigma) - w^\varepsilon(\sigma)) \sum_{(x,e)\subset\sigma} g^\varepsilon(x,e,m^\varepsilon(x,e)). \end{aligned}$$

where $G^{\varepsilon\prime}(x,e,\cdot) = g^\varepsilon(x,e,\cdot)$ is the derivative of $G^\varepsilon(x,e,\cdot)$. Then

$$\sum_{\sigma\in C^\varepsilon} w^\varepsilon(\sigma)\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma) \leq \sum_{\sigma\in C^\varepsilon} v(\sigma)\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma).$$

By minimizing the right-hand side, we obtain

$$\sum_{(x,y)\in N^\varepsilon\times N^\varepsilon} \sum_{\sigma\in C^\varepsilon_{x,y}} w^\varepsilon(\sigma)\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma) = \sum_{(x,y)\in N^\varepsilon\times N^\varepsilon} \gamma^\varepsilon(x,y) \min_{\sigma'\in C^\varepsilon_{x,y}} \tau^\varepsilon_{\mathbf{m}^\varepsilon}(\sigma')$$

which exactly is the definition of Wardrop equilibrium. To prove the converse, it is sufficient to notice this is a convex problem (since the functions g^ε are nondecreasing with respect to the last variable) so that the inequality above is enough for a global minimum. □

As mentioned in the proof, the problem (2.4) is convex so we can easily obtain existence results and numerical schemes. Unfortunately, this problem becomes quickly costly whenever the network is very dense, since it requires to enumerate all paths flows $w^\varepsilon(\sigma)$. For this reason, we can not use this characterization to study realistic congested networks. An alternative consists in working with the dual formulation which is

$$\inf_{t^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}} \sum_{(x,e)\in E^\varepsilon} H^\varepsilon(x,e,t^\varepsilon(x,e)) - \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y)T^\varepsilon_{\mathbf{t}^\varepsilon}(x,y), \tag{2.5}$$

where $\mathbf{t}^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}$ should be understood as $\mathbf{t}^\varepsilon = (t^\varepsilon(x,e))_{(x,e)\in E^\varepsilon}$, $H^\varepsilon(x,e,.) := (G^\varepsilon(x,e,.))^*$ is the Legendre transform of $G^\varepsilon(x,e,.)$ that is

$$H^\varepsilon(x,e,t) := \sup_{m\geq 0}\{mt - G^\varepsilon(x,e,m)\},\ \forall t\in\mathbb{R}_+ \tag{2.6}$$

and $T^\varepsilon_{\mathbf{t}^\varepsilon}$ is the minimal length functional:

$$T^\varepsilon_{\mathbf{t}^\varepsilon}(x,y) = \min_{\sigma\in C^\varepsilon_{x,y}} \sum_{(z,e)\subset\sigma} t^\varepsilon(z,e).$$

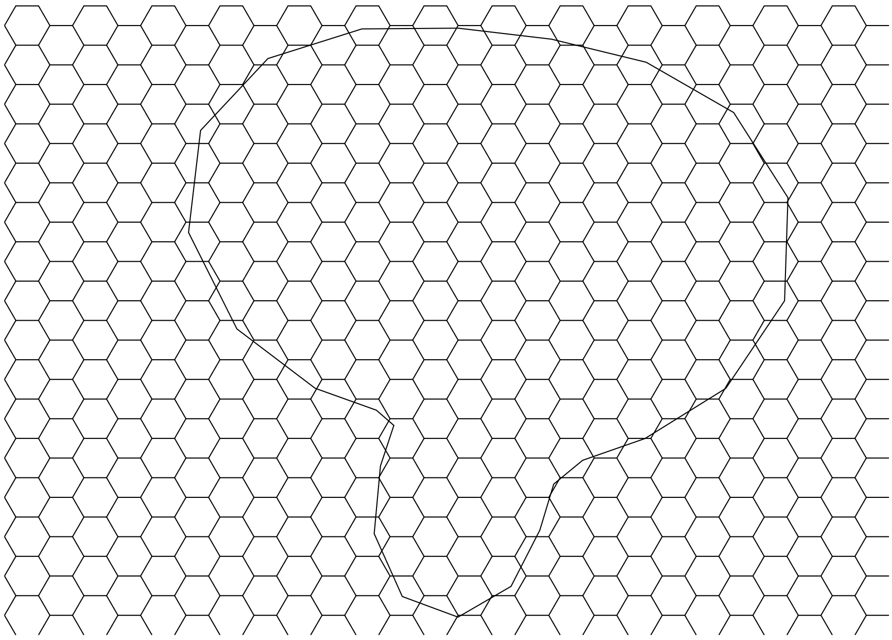

Figure 2.1 – An example of domain in $2d$-hexagonal model

The complexity of (2.5) seems better since we only have $\#E^\varepsilon = O(\varepsilon^{-d})$ nodes variables. However an important disadvantage appears in the dual formulation. The term $T^\varepsilon_{\mathbf{t}^\varepsilon}$ is nonsmooth, nonlocal and we might have difficulties to optimize that. Nevertheless we will see that we may pass to a continuous limit which will simplify the structure since we can then use the Hamilton-Jacobi theory.

3 The Γ-convergence result

3.1 Assumptions

We obviously have to make some structural assumptions on the ε-dependence of the networks and the data to be able to pass to a continuous limit in the Wardrop equilibrium problem. To understand all these assumptions, we will illustrate with some examples. Here we will consider the cases of regular decomposition (cartesian, triangular and hexagonal, see Figure 2.1) for $d = 2$. In these models, all the arcs in Ω_ε have the same length that is ε. We will introduce some notations and to refer to a specific example, we will simply add the letters c (for the cartesian case), t (for the triangular one) and h (for the hexagonal one).

The first assumption concerns the length of the arcs in the networks.

Assumption 2.1. *There exists a constant $C > 0$ such that for every $\varepsilon > 0, (x, e) \in E^\varepsilon$, we have $C\varepsilon \leq |e| \leq \varepsilon$.*

More generally we denote by C a generic constant that does not depend on the scale parameter ε.

The following assumption is on the discrete network Ω_ε, $\varepsilon > 0$. Roughly speaking, the arcs of Ω_ε define a bounded polyhedron (still denoted by Ω_ε by abuse of notations) that is an approximation of $\overline{\Omega}$. The set Ω_ε is the union of cells - or polytopes - (V_i^ε). Each V_i^ε is itself the union of subcells - or facets - $(F_{i,k}^\varepsilon)$. More precisely, we have:

Assumption 2.2. *Up to a subsequence, $\Omega_\varepsilon \subset \Omega_{\varepsilon'}$ for $\varepsilon > \varepsilon' > 0$ and $\overline{\Omega} = \bigcup_{\varepsilon>0} \Omega_\varepsilon$. For $\varepsilon > 0$, we have*

$$\Omega_\varepsilon = \bigcup_{i \in I^\varepsilon} V_i^\varepsilon, \ I^\varepsilon \text{ is finite.}$$

There exists $S \in \mathbb{N}$ and $s_i \leq S$ such that for every $\varepsilon > 0$ and $i \in I^\varepsilon$, $V_i^\varepsilon = Conv(x_{i,1}^\varepsilon, \dots, x_{i,s_i}^\varepsilon)$ where for $j = 1, \dots, s_i - 1$, $x_{i,j}^\varepsilon$ is a neighbor of $x_{i,j+1}^\varepsilon$ and x_{i,s_i}^ε is a neighbor of $x_{i,1}^\varepsilon$ in Ω_ε. Its interior $\mathring{V}_i^\varepsilon$ contains no arc $\in E^\varepsilon$. For $i \neq j, V_i^\varepsilon \cap V_j^\varepsilon = \emptyset$ or exactly a facet (of dimension $\leq d-1$). Let us denote X_i^ε the isobarycenter of all nodes $x_{i,k}^\varepsilon$ contained in V_i^ε. For $i \in I^\varepsilon$, we have

$$V_i^\varepsilon = \bigcup_{j \in I_i^\varepsilon} F_{i,j}^\varepsilon \text{ with } I_i^\varepsilon \text{ finite.}$$

For $j \in I_i^\varepsilon$, $F_{i,j}^\varepsilon = Conv(X_i^\varepsilon, x_{i,j_1}^\varepsilon, \dots, x_{i,j_d}^\varepsilon)$ with these $(d+1)$ points that are affinely independent. For every $k \neq l$, $F_{i,k}^\varepsilon \cap F_{i,l}^\varepsilon$ is a facet of dimension $\leq d-1$, containing X_i^ε. There exists a constant $C > 0$ independent of ε such that for every $i \in I^\varepsilon$ and $j \in I_i^\varepsilon$, the volume of V_i^ε and $F_{i,j}^\varepsilon$ satisfies

$$\frac{1}{C}\varepsilon^d < |F_{i,j}^\varepsilon| \leq |V_i^\varepsilon| < C\varepsilon^d.$$

The estimate on the volumes implies another estimate: for every $i \in I^\varepsilon$ and $k = 1, \dots, s_i$, we have

$$\frac{1}{C}\varepsilon < \operatorname{dist}(X_i^\varepsilon, x_{i,k}^\varepsilon) < C\varepsilon \tag{2.7}$$

(the constant C is not necessarily the same). This hypothesis will in particular allow us to make possible a discretization of Morrey's theorem and then to prove Lemma 2.6. It is a non trivial extension of what is happening in dimension 2. Figure 2.2 shows an illustration of Assumption 2.2 in the cartesian case.

We must also impose some technical assumptions on N^ε and E^ε.

Assumption 2.3. *There exists $N \in \mathbb{N}, D = \{v_k\}_{k=1,\dots,N} \in C^{0,\alpha}(\mathbb{R}^d, \mathbb{S}^{d-1})^N$ and $\{c_k\}_{k=1,\dots,N} \in C^1(\overline{\Omega}, \mathbb{R}_+^*)^N$ with $\alpha > d/p$ such that E^ε weakly converges in the sense that*

$$\lim_{\varepsilon \to 0^+} \sum_{(x,e) \in E^\varepsilon} |e|^d \varphi\left(x, \frac{e}{|e|}\right) = \int_{\Omega \times \mathbb{S}^{d-1}} \varphi(x,v)\, \theta(dx,dv), \forall \varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}),$$

where $\theta \in \mathcal{M}_+(\Omega \times \mathbb{S}^{d-1})$ is of the form

$$\theta(dx,dv) = \sum_{k=1}^N c_k(x) \delta_{v_k(x)} dx.$$

The v_k's are the possible directions in the continuous model. We have to keep in mind that for every $x \in \mathbb{R}^d$, the $v_k(x)$ are not necessarily pairwise distinct. The requirement $v_k \in C^{0,\alpha}(\mathbb{R}^d)$ with $\alpha > d/p$ is technical and will in particular be useful

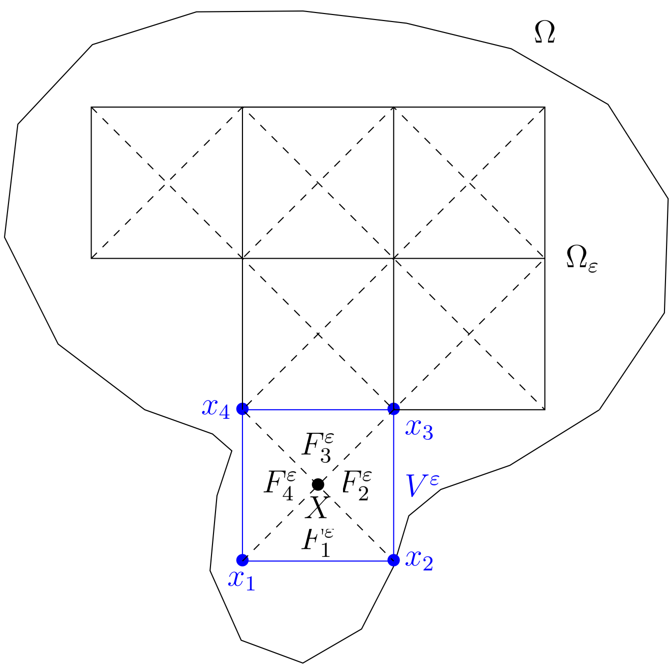

Figure 2.2 – An illustration of Assumption 2.2 in the cartesian case for $d = 2$.

to prove Lemma 2.7. In our examples, the sets of directions are constant and we have

$$D_c = \{v_{c_1} = (1,0), v_{c_2} = (0,1), v_{c_3} = (-1,0), v_{c_4} = (0,-1)\},$$
$$D_t = D_h = \{v_{t_k} = e^{i\pi/6} \cdot e^{i\pi(k-1)/3}\}_{k=1,\dots,6}.$$

The c_k's are the volume coefficients. In our examples, they are constant and do not depend on k:

$$c_c = 1, c_t = \frac{2}{\sqrt{3}} \text{ and } c_h = \frac{2}{3\sqrt{3}}.$$

We notice that the c_l's are different. Indeed, a square whose side length is ε does not have the same area as a hexagon whose side length is ε. The next assumption imposes another condition on the directions v_k's.

Assumption 2.4. *There exists a constant $C > 0$ such that for every $(x, z, \xi) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \times \mathbb{R}_+^N$, there exists $\bar{Z} \in \mathbb{R}_+^N$ such that $|\bar{Z}| \leq C$ and*

$$\bar{Z} \cdot \xi = \min\{Z \cdot \xi; Z = (z_1, \dots, z_N) \in \mathbb{R}_+^N \textit{ and } \sum_{k=1}^N z_k v_k(x) = z\}. \tag{2.8}$$

This means that the family $\{v_k\}$ is positively generating and that for every $(x, z) \in \mathbb{R}^d \times \mathbb{R}^d$, a conical decomposition of z, not too large compared to z, is always possible in the family $D(x)$. In the cartesian case, for all $z \in \mathbb{R}^2$, we have an unique (interesting) conical decomposition of z in D_c while in the other examples, we always have the existence but not the uniqueness. The existence of a controlled minimizer allows us to keep some control over "the" decomposition of $z \in \mathbb{R}^d$ in the family of directions. We now see a counterexample that looks like the cartesian case. We take $N = 4, d = 2, v_1 = (1,0), v_3 = (-1,0)$ and $v_4 = (0,1)$ (these directions are constant). We assume that there exists x_0 in Ω such that $v_2(x) \to -v_4 = (0,1)$ as $x \to x_0$ and $v_{21}(x) > 0$ for $x \neq x_0$ where $v_2(x) = (v_{21}(x), v_{22}(x))$. Then for every $z = (z_1, z_2) \in \mathbb{R}^2$ such that $z_1 > 0$, we can write $z = \lambda_2(x)v_2(x) + \lambda_4(x)v_4(x)$ for x close enough to x_0 with

$$\lambda_2(x) = \frac{z_1}{v_{21}(x)} \text{ and } \lambda_4(x) = z_1\frac{v_{22}(x)}{v_{21}(x)} - z_2.$$

Then $\lambda_2(x)$ and $\lambda_4(x) \to +\infty$ as $x \to x_0$. For $\xi = (1,0,1,0)$, the value of (4.6) always is zero but the only decomposition that solves the problem is not controlled. This example shows that the existence of a controlled decomposition does not imply that the minimal decomposition is controlled. However this assumption is still natural since we want to control the flow on each arc in order to minimize the transport cost.

Assumption 2.5. *Up to a subsequence, E^ε may be partitioned as $E^\varepsilon = \bigsqcup_{k=1}^N E_k^\varepsilon$ such that for every $k = 1, \dots, N$, one has*

$$\lim_{\varepsilon \to 0^+} \sum_{(x,e)\in E_k^\varepsilon} |e|^d \varphi\left(x, \frac{e}{|e|}\right) = \int_\Omega c_k(x)\varphi(x, v_k(x))\,dx,\ \forall \varphi \in C(\Omega \times \mathbb{S}^{d-1}, \mathbb{R}). \quad (2.9)$$

and for every $(x,e) \in E_k^\varepsilon$,

$$\left|\frac{e}{|e|} - v_k(x)\right| = O(1). \quad (2.10)$$

This hypothesis is natural. Indeed, for every $x \in \Omega$ the condition $c_k(x) > 0$ implies that for $\varepsilon > 0$ small enough, there exists an arc (y,e) in E^ε such that y is near x and $e/|e|$ close to $v_k(x)$. We will use it particularly to prove Lemma 2.8. The next assumption is more technical and will specifically serve to apply Lemma 2.7.

Assumption 2.6. *For $\varepsilon > 0$, there exist d (finite) sets of paths $C_1^\varepsilon, \dots, C_d^\varepsilon$ and d linearly independent functions $e_1, \dots, e_d : \Omega \to \mathbb{R}^d$ such that for every $x \in \Omega, i = 1, \dots, d$, $e_i(x) = \sum_k \alpha_k^i c_k(x)v_k(x)$ where for $k = 1, \dots, N$, α_k^i is constant and equal to 0 or 1 so that for $i = 1, \dots, d$, we have*

$$\bigcup_{\sigma \in C_i^\varepsilon} \{(x,e) \subset \sigma\} = \{(x,e) \in E^\varepsilon / \exists k \in \{1, \dots, N\}, \alpha_k^i = 1 \textit{ and } (x,e) \in E_k^\varepsilon\}.$$

For every $(\sigma, \sigma') \in C_i^\varepsilon \times C_i^\varepsilon$, if $\sigma \neq \sigma'$ then $\sigma \cap \sigma' = \emptyset$. We assume

$$\max_{\sigma = (y_0, \dots, y_{N(\sigma)}) \in C_i^\varepsilon} (dist(y_0, \partial\Omega), dist(y_{N(\sigma)}, \partial\Omega)) \to 0 \textit{ as } \varepsilon \to 0.$$

and

$$||y_{k+1} - y_k|^{d-1} - |y_k - y_{k-1}|^{d-1}| = O(\varepsilon^d)$$

for $\sigma = (y_0, \dots, y_{N(\sigma)}) \in C_i^\varepsilon$ and $k = 2, \dots, N(\sigma) - 1$.

Formally speaking, it means that we partition points into a set of disjoint paths whose extremities tend to the boundary of Ω and such that we may do a change of variables for the derivates. For φ a regular function, $\nabla\varphi$ then becomes

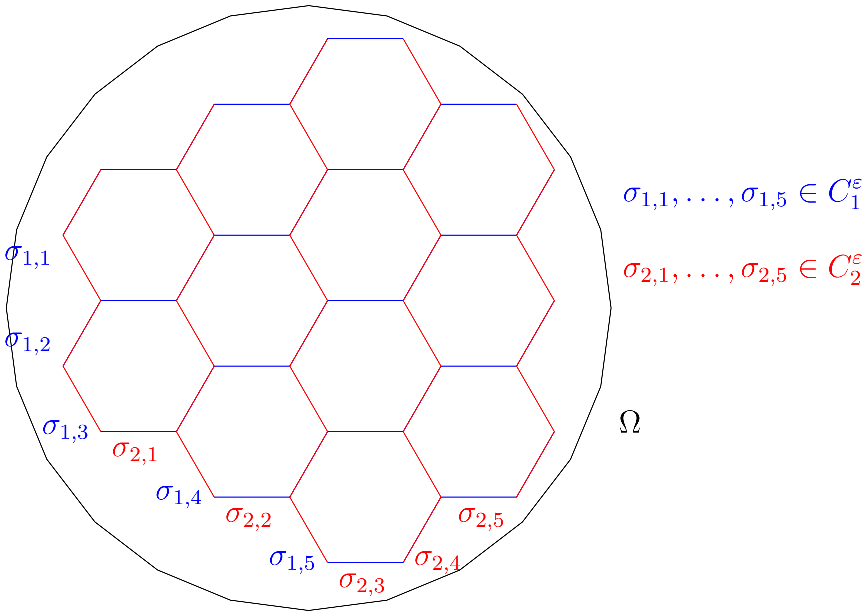

Figure 2.3 – An illustration of Assumption 2.6 in the hexagonal case for $d = 2$

$(\partial_{e_1}\varphi, \dots, \partial_{e_d}\varphi)$. Note that the condition on the modules is a sort of volume elements in spherical coordinates, dV being approximately $r^{d-1}dr$. In our examples, this requirement is trivial since arcs length is always equal to ε in Ω_ε. It will allow us to prove some statements on functions. More specifically, we will need to show that some functions have (Sobolev) regularity properties. We can then apply Lemma 2.7. For this purpose, it will be simpler to use these e_i. In the cartesian case ($d = 2$), we can simply take $e_{c_1} = v_{c_1}$ and $e_{c_2} = v_{c_2}$ and in the triangular one, we can write $e_{t_1} = c_t v_{t_1}$ and $e_{t_2} = c_t v_{t_2}$. In the hexagonal case, we can for instance define $e_{h_1} = c_h(v_{h_1} + v_{h_2})$ and $e_{h_2} = c_h(v_{h_2} + v_{h_3})$. This last example is illustrated by Figure 2.3 (with Ω being a circle). We can notice that some arcs are in paths $\sigma \in C_1^\varepsilon$ and $\sigma' \in C_2^\varepsilon$.

In particular, thanks to Assumption 2.6, for all $k = 1, \dots, N$, there exists d continuous functions $\lambda_1^k, \dots, \lambda_d^k : \Omega \mapsto \mathbb{R}$ such that for all $x \in \Omega$, we have

$$c_k(x)v_k(x) = \sum_{i=1}^{d} \lambda_i^k(x)e_i(x) = \sum_{l=1}^{N} \mu_l^k(x)c_k(x)v_k(x).$$

where the functions μ_l^k are given by the relation $\mu_l^k = \sum_{i=1}^N \alpha_l^i \lambda_i^k$. The following assumption is on these functions μ_l^k and will be crucial to prove Lemma 2.8.

Assumption 2.7. *For all $k = 1, \dots, N$, there exist integers $n_k, m_1^k, \dots, m_N^k$ ($n_k > 0$) such that for every $x \in \Omega$ we have*

$$n_k c_k(x)v_k(x) = \sum_{l=1}^{N} m_l^k c_l(x)v_l(x).$$

We denote $N_k = \sum_{l=1}^N m_l^k$. For $\varepsilon > 0$, for every $(x, e) \in E_k^\varepsilon$, there exists a path

$\sigma_{(x,e)} \in C^\varepsilon, \sigma_{(x,e)} = (x_0, \dots, x_{N_k})$ *such that*

$$|x_0 - x| = O(1),$$
$$||x_{N_k} - x_0|^{d-1} - n_k|e|^{d-1}| = O(\varepsilon^d),$$
$$||x_{l+1} - x_l|^{d-1} - |x_l - x_{l-1}|^{d-1}| = O(\varepsilon^d) \ \textit{for } l = 2, \dots, N_k - 1$$

and

$$\sigma_{(x,e)} = \bigcup_{l=1}^{d} \{m_l^k \textit{ arcs in } E_l^\varepsilon\}.$$

It means that in the discrete network for every arc $(x,e) \in E_k^\varepsilon$, there is a path around x whose arcs have almost the same length and such that the distance between the origin and the destination is equal to $n_k|e|$. It will allow us to rewrite an union of districting paths in a simpler sum.

All these structural hypothesis on Ω^ε are satisfied in our three classical examples. The cartesian one is the most obvious and the hexagonal one is a subcase of the triangular one. The following assumption is on the transport plan.

Assumption 2.8. *The family of discrete functions* $(\varepsilon^{\frac{d}{2}-1}\gamma^\varepsilon)_{\varepsilon>0}$ *weakly star converges to a finite nonnegative measure* γ *on* $\overline{\Omega} \times \overline{\Omega}$ *in the sense that the family of discrete measures*
$\varepsilon^{\frac{d}{2}-1} \sum_{(x,y)\in N^\varepsilon \times N^\varepsilon} \gamma^\varepsilon(x,y)\delta_{(x,y)}$ *weakly star converges to* γ*:*

$$\lim_{\varepsilon\to 0^+} \varepsilon^{\frac{d}{2}-1} \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y)\varphi(x,y) = \int_{\overline{\Omega}\times\overline{\Omega}} \varphi \, d\gamma; \ \forall \varphi \in C(\overline{\Omega} \times \overline{\Omega}). \tag{2.11}$$

The next assumption focuses on the congestion functions g^ε.

Assumption 2.9. *The function* g^ε *is of the form*

$$g^\varepsilon(x,e,m) = |e|^{d/2} g\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d/2}}\right), \ \forall \varepsilon > 0, (x,e) \in E^\varepsilon, m \geq 0 \tag{2.12}$$

where $g : \Omega \times \mathbb{S}^{d-1} \times \mathbb{R}_+ \mapsto \mathbb{R}$ *is a given continuous, nonnegative function that is increasing in its last variable.*

Recalling (2.4) and (2.6) we then have with Assumption 2.9 that

$$G^\varepsilon(x,e,m) = |e|^d G\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d/2}}\right) \text{ where } G(x,v,m) := \int_0^m g(x,v,\alpha)d\alpha$$

and

$$H^\varepsilon(x,e,t) = |e|^d H\left(x, \frac{e}{|e|}, \frac{t}{|e|^{d/2}}\right) \text{ where } H(x,v,\cdot) := (G(x,v,\cdot))^\star$$

i.e. for every $\xi \in \mathbb{R}_+$:

$$H\left(x, \frac{e}{|e|}, \xi\right) := \sup_{m\in\mathbb{R}_+} \left\{ m\xi - G\left(x, \frac{e}{|e|}, m\right)\right\}.$$

For every $(x,v) \in \overline{\Omega} \times \mathbb{S}^{d-1}$, $H(x,v,\cdot)$ is actually strictly convex since $G(x,v,\cdot)$ is C^1 (thanks to Assumption 2.9). Note that in the case $d=2$, we have

$$g^\varepsilon(x,e,m) = |e| g\left(x, e, \frac{m}{|e|}\right) \text{ for all } \varepsilon > 0, (x,e) \in E^\varepsilon, m \geq 0.$$

It is natural. It means that the traveling time on an arc of length $|e|$ is of order $|e|$ and depends on the flow per unit of length i.e. $m/|e|$. For the general case, we have extended this assumption. The exponent $d/2$ is not very natural, it does not represent a physical phenomenon but it allows us to obtain the same relation between G and H, that means, $H(x,e,\cdot)$ is the Legendre transform of $G(x,e,\cdot)$. Moreover, we may approach some integrals by sums. We can also note that there is actually no ε-dependence on the g^ε.

We also add assumptions on H:

Assumption 2.10. *H is continuous with respect to the first two arguments and there exists $p > d$ and two constants $0 < \lambda \leq \Lambda$ such that for every $(x,v,\xi) \in \overline{\Omega} \times \mathbb{S}^{d-1} \times \mathbb{R}_+$ one has*

$$\lambda(\xi^p - 1) \leq H(x,v,\xi) \leq \Lambda(\xi^p + 1). \tag{2.13}$$

The p-growth is natural since we want to work in L^p in the continuous limit. The condition $p > d$ has a technical reason, that will allow us to use Morrey's inequality. That will be crucial to pass to the limit in the nonlocal term that contains $T^\varepsilon_{\mathbf{t}^\varepsilon}$.

3.2 The limit functional

In view of the previous paragraph and in particular Assumption 2.9, it is natural to rescale the arc-times $\mathbf{t}^\varepsilon$ by defining new variables

$$\xi^\varepsilon(x,e) := \frac{t^\varepsilon(x,e)}{|e|^{d/2}} \text{ for all } (x,e) \in E^\varepsilon, \tag{2.14}$$

i.e. for every $(x,e) \in E^\varepsilon$,

$$\xi^\varepsilon(x,e) = \frac{g^\varepsilon(x,e,m^\varepsilon(x,e))}{|e|^{d/2}} = g\left(x, \frac{e}{|e|}, \frac{m^\varepsilon(x,e)}{|e|^{d/2}}\right).$$

Remark 2.1. *In fact, we could choose another exponent. The term $d/2$ allows to have the same exponent in* (2.12) *and* (2.14)*. More generally, instead of mainly* (Assumption 2.8)*,* (Assumption 2.9) *and* (2.14) *we could write with $\beta \in \mathbb{R}$:*

1. *The sequence $(\varepsilon^{\beta-1}\gamma^\varepsilon)_{\varepsilon>0}$ weakly star converges to a finite nonnegative measure γ on $\overline{\Omega} \times \overline{\Omega}$ in the sense that the family of discrete measures $\varepsilon^{\beta-1} \sum_{(x,y)\in N^\varepsilon \times N^\varepsilon} \gamma^\varepsilon(x,y)\delta_{(x,y)}$ weakly star converges to γ:*

$$\lim_{\varepsilon \to 0^+} \varepsilon^{\beta-1} \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y)\varphi(x,y) = \int_{\overline{\Omega}\times\overline{\Omega}} \varphi \, d\gamma; \;\; \forall \varphi \in C(\overline{\Omega} \times \overline{\Omega}). \tag{2.15}$$

2. *The function g^ε is of the form*

$$g^\varepsilon(x,e,m) = |e|^\beta g\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d-\beta}}\right), \; \forall \varepsilon > 0, (x,e) \in E^\varepsilon, m \geq 0 \tag{2.16}$$

with the same requirements on g.

3.

$$\xi^\varepsilon(x,e) = \frac{t^\varepsilon(x,e)}{|e|^\beta} \text{ for all } (x,e) \in E^\varepsilon. \tag{2.17}$$

Here we prefer to take a special β for the sake of clarity but the reasoning is the same in the next sections with another exponent.

Then rewrite the formula (2.5) in terms of ξ^ε as:

$$\inf_{\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}} J^\varepsilon(\xi^\varepsilon) := I_0^\varepsilon(\xi^\varepsilon) - I_1^\varepsilon(\xi^\varepsilon) \tag{2.18}$$

where

$$I_0^\varepsilon(\xi^\varepsilon) := \sum_{(x,e)\in E^\varepsilon} |e|^d H\left(x, \frac{e}{|e|}, \xi^\varepsilon(x,e)\right) \tag{2.19}$$

and

$$I_1^\varepsilon(\xi^\varepsilon) := \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x,y) \left(\min_{\sigma\in C_{x,y}^\varepsilon} \sum_{(z,e)\subset\sigma} |e|^{d/2}\xi^\varepsilon(z,e)\right). \tag{2.20}$$

In view of Assumption 2.10, let us denote

$$L_+^p(\theta) := \{\xi \in L^p(\Omega \times \mathbb{S}^{d-1}, \theta), \xi \geq 0\}.$$

It is natural to introduce

$$I_0(\xi) := \int_{\Omega\times\mathbb{S}^{d-1}} H(x,v,\xi(x,v))\theta(dx,dv),\ \forall \xi \in L_+^p(\theta), \tag{2.21}$$

as the continuous limit of I_0^ε. It is more involved to find the term that plays the same role as I_1^ε since we must define some integrals on paths. First rearrange the second term (2.20) as an integral. Let $\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}, (x,y) \in N^\varepsilon \times N^\varepsilon$ and $\sigma \in C_{x,y}^\varepsilon$, $\sigma = (x_0, ..., x_{N(\sigma)})$. Let us extend σ on $[0, N(\sigma)]$ in a piecewise affine function by defining $\sigma(t) = \sigma(k) + (t-k)(\sigma(k+1) - \sigma(k))$ and ξ^ε in a piecewise constant function : $\xi^\varepsilon(\sigma(t), \dot\sigma(t)) = \xi^\varepsilon(\sigma(k), \sigma(k+1) - \sigma(k))$ for $t \in [k, k+1]$. For every $(x,e) \in E^\varepsilon$ we call $\Psi^\varepsilon(x,e)$ the "canonical" decomposition of e on $D(x)$. More precisely, recalling Assumption 2.5, for $(x,e) \in E^\varepsilon$, there exists $k_{(x,e)} \in \{1, \ldots, N\}$ such that $(x,e) \in E_{k_{(x,e)}}^\varepsilon$ and then we set

$$\Psi^\varepsilon(x,e) = (0, \ldots, \underbrace{|e|}_{k_{(x,e)}\text{th coordinate}}, \ldots, 0) \in \mathbb{R}^N.$$

For $\sigma \in C^\varepsilon$ and $t \in [k, k+1[$, we write $\Psi^\varepsilon(\sigma(t), \dot\sigma(t)) = \Psi^\varepsilon(\sigma(k), \sigma(k+1) - \sigma(k))$. Let us also define a function $\boldsymbol{\xi}^\varepsilon$ as follows:

$$\boldsymbol{\xi}^\varepsilon(x) := \sum_{e/(x,e)\in E^\varepsilon} \frac{\xi^\varepsilon(x,e)}{|e|} \Psi^\varepsilon(x,e) \text{ for } x \in N^\varepsilon.$$

Let us extend $\boldsymbol{\xi}^\varepsilon$ in a piecewise constant function on the arcs of E^ε: let $y \in \Omega$ such that there exists $(x,e) \in E^\varepsilon$ with $y \in (x,e)$, then we define

$$\boldsymbol{\xi}^\varepsilon(y) := \sum_{(x,e)\in E^\varepsilon/y\in(x,e)} \frac{\xi^\varepsilon(x,e)}{|e|} \Psi^\varepsilon(x,e).$$

This definition is consistent since the arcs that appear in the sum are in some different E_k^ε. By abuse of notations, we continue to write σ, ξ^ε and $\boldsymbol{\xi}^\varepsilon$ for these new functions. Thus we have

$$\begin{aligned}\sum_{(x,e)\subset\sigma} |e|\xi^\varepsilon(x,e) &= \sum_{k=0}^{N(\sigma)-1} |\sigma(k+1)-\sigma(k)|\xi^\varepsilon(\sigma(k),\sigma(k+1)-\sigma(k))\\ &= \sum_{k=0}^{N(\sigma)-1} \Psi^\varepsilon(\sigma(k),\sigma(k+1)-\sigma(k))\cdot\boldsymbol{\xi}^\varepsilon(\sigma(k))\\ &= \int_0^{N(\sigma)} \Psi^\varepsilon(\sigma(t),\dot\sigma(t))\cdot\boldsymbol{\xi}^\varepsilon(\sigma(t))dt.\end{aligned}$$

We then get

$$\begin{aligned}\min_{\sigma\in C^\varepsilon_{x,y}} \sum_{(x,e)\subset\sigma} |e|\xi^\varepsilon(x,e) &= \inf_{\sigma\in C^\varepsilon_{x,y}} \int_0^{N(\sigma)} \Psi^\varepsilon(\sigma(t),\dot\sigma(t))\cdot\boldsymbol{\xi}^\varepsilon(\sigma(t))dt\\ &= \inf_{\sigma\in C^\varepsilon_{x,y}} \int_0^1 \Psi^\varepsilon(\tilde\sigma(t),\dot{\tilde\sigma}(t))\cdot\boldsymbol{\xi}^\varepsilon(\tilde\sigma(t))dt\end{aligned}$$

where $\tilde\sigma : [0,1] \to \overline{\Omega}$ is the reparameterization $\tilde\sigma(t) = \sigma(N(\sigma)t), t \in [0,1]$. For every $x \in \overline{\Omega}, z \in \mathbb{R}^d$ let us define

$$A_x^z := \left\{ Z \in \mathbb{R}_+^N, Z = (z_1,\dots,z_N) / \sum_{k=1}^N z_k v_k(x) = z \right\}.$$

Then for every $\xi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R}_+)$, define

$$\begin{aligned}c_\xi(x,y) &:= \inf_{\sigma\in C_{x,y}} \left\{ \int_0^1 \inf_{(x_1,\dots,x_N)\in A_{\sigma(t)}^{\dot\sigma(t)}} \left(\sum_{k=1}^N x_k \xi(\sigma(t), v_k(\sigma(t))) \right) dt \right\}\\ &= \inf_{\sigma\in C_{x,y}} \inf_{\rho\in\mathcal{P}_\sigma} \int_0^1 \left(\sum_{k=1}^N \xi(\sigma(t), v_k(\sigma(t)))\, \rho_k(t) \right) dt\\ &= \inf_{\sigma\in C_{x,y}} \inf_{\rho\in\mathcal{P}_\sigma} L_\xi(\sigma,\rho),\end{aligned}$$

where

$$\mathcal{P}_\sigma := \left\{ \rho : t\in[0,1] \to \rho(t) \in \mathbb{R}_+^N / \dot\sigma(t) = \sum_{k=1}^N v_k(\sigma(t))\,\rho_k(t) \text{ a.e. } t \right\},$$

$$L_\xi(\sigma,\rho) := \int_0^1 \left(\sum_{k=1}^N \xi(\sigma(t), v_k(\sigma(t)))\, \rho_k(t) \right) dt$$

and $C_{x,y}$ is the set of absolutely continuous curves σ with values in $\overline{\Omega}$ and such that $\sigma(0) = x$ and $\sigma(1) = y$. For every $x \in \Omega$ and $z \in \mathbb{R}^d$, there is no a priori uniqueness of the decomposition of z in the family $\{v_k(x)\}_k$ so that we have to take the infimum over all possible decompositions. The definition of $\mathcal{P}_\sigma$ takes in account this constraint. For every $\rho \in \mathcal{P}_\sigma$ and $t \in [0,1]$ the terms $\rho_k(t)$ are the weights of $\dot\sigma(t)$ in the family $\{v_k(\sigma(t)\}$. (σ,ρ) is a sort of generalized curve. It will allow us to distinguish between different limiting behaviors. A simple example is the following approximations, $\varepsilon > 0$:

In both cases It is the same curve σ: the straight line from x to y. But the ρ are different. Indeed, in the left case, we have only one direction that is $(1,0)$, it is a line. In the right one, we have two directions that tend to $(0,1)$ and $(0,-1)$ as ε tends to 0, these are oscillations.

Now, our aim is to extend the definition of c_ξ to the case where ξ is only $L^p_+(\theta)$. We will strongly generalize the method used in [39] to the case of generalized curves. First, let us notice that we may write c_ξ in another form:

$$c_\xi(x,y) = \inf_{\sigma \in C_{x,y}} \widetilde{L}_\xi(\sigma) \ \text{ for } \xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+),$$

where

$$\widetilde{L}_\xi(\sigma) := \int_0^1 \Phi_\xi(\sigma(t), \dot{\sigma}(t))\, dt \tag{2.22}$$

with for all $x \in \overline{\Omega}$ and $y \in \mathbb{R}^d$, $\Phi_\xi(x,y)$ being defined as follows:

$$\begin{aligned}\Phi_\xi(x,y) &:= \inf_{Y \in \mathbb{R}^N_+} \left\{ \sum_{k=1}^N y_k \xi(x, v_k(x)) : Y = (y_1, \dots, y_N) \in A^y_x \right\} \\ &= \inf_{Y \in A^y_x} Y \cdot \boldsymbol{\xi}(x),\end{aligned}$$

where $\boldsymbol{\xi}(x) := (\xi(x, v_1(x)), \dots, \xi(x, v_N(x)))$.

The next lemma shows that Φ_ξ defines a sort of Finsler metric. It is an anisotropic model but Φ_ξ is not even and so c_ξ is not symmetric. Moreover c_ξ is not necessarily strictly positive between two different points so that c_ξ is not a distance. However $\Phi_\xi(x,\cdot)$ looks like a norm that depends on the point $x \in \overline{\Omega}$. Its unit ball is a polyhedron in $\mathbb{R}^d$ that changes with x. Formally speaking, the minimizing element $Y = (y_1, \dots, y_N)$ represents the coefficients for the "Finsler distance" c_ξ.

Lemma 2.1. *Let $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$. Then one has:*

1. *The infimum in the function Φ_ξ is attained and Φ_ξ is continuous.*
2. *For all $x \in \overline{\Omega}$, $\Phi_\xi(x,\cdot)$ is homogeneous of degree 1 and convex.*
3. *If $\{\xi_n\}_n$ is a sequence in $C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ and $x \in \overline{\Omega}$ such that $\{\boldsymbol{\xi_n}(x)\}_n$ converges to $\boldsymbol{\xi}(x)$ then $\{\Phi_{\xi_n}(x,y)\}_n$ converges to $\Phi_\xi(x,y)$ for all $y \in \mathbb{R}^d$.*

Proof. Let $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$.

1.) First notice that Φ_ξ actually is a minimum thanks to Assumption 2.4. Let $(x,y) \in \overline{\Omega} \times \mathbb{R}^d$ and sequences $\{x^n, y^n\}_{n \geq 0} \in (\overline{\Omega} \times \mathbb{R}^d)^{\mathbb{N}}$ converging to (x,y). Still thanks to Assumption 2.4, there exists $Y_n \in A^{y^n}_{x^n}$ and $Y' \in A^y_x$ such that $|Y_n| \leq C|y^n|$, $\Phi_\xi(x^n, y^n) = Y_n \cdot \boldsymbol{\xi}(x^n)$ and $\Phi_\xi(x,y) = Y' \cdot \boldsymbol{\xi}(x)$. For n large enough, we have $|Y_n| \leq 2C|y|$ so that up to a subsequence, there exists $Y \in A^y_x$ such that Y_n converges to Y. Then we have

$$\Phi_\xi(x^n, y^n) = Y_n \cdot \boldsymbol{\xi}(x^n) \underset{n \to +\infty}{\longrightarrow} Y \cdot \boldsymbol{\xi}(x) \geq \Phi_\xi(x,y).$$

Now we build a sequence $Z_n \in A_{x^n}^{y^n}$ converging to Y'. Set $z^n = y^n - \sum_{k=1}^N y'_k v_k(x^n)$, where $Y' = (y'_1, \dots, y'_N)$. Let $\varepsilon^n \in A_{x^n}^{z^n}$ such that $|\varepsilon^n| \leq C|z^n|$. Then define $Z_n = Y' + \varepsilon^n \in A_{x^n}^{y^n}$. Since $z^n \to 0$ as $n \to +\infty$, Z_n converges to Y'. Passing to the limit in the inequality $Z_n \cdot \boldsymbol{\xi}(x^n) \geq \Phi_\xi(x^n, y^n)$, we obtain $Y' \cdot \boldsymbol{\xi}(x) \geq Y \cdot \boldsymbol{\xi}(x)$. That is, we have $Y \cdot \boldsymbol{\xi}(x) = \Phi_\xi(x, y)$ and so Φ_ξ is continuous.

2.) The statement of homogeneity is obvious. Let $(x, y_1, y_2) \in \overline{\Omega} \times \mathbb{R}^d \times \mathbb{R}^d, t \in [0,1], Y_1 \in A_x^{y_1}$ and $Y_2 \in A_x^{y_2}$ such that $\Phi_\xi(x, y_i) = Y_i \cdot \xi(x)$. Then $tY_1 + (1-t)Y_2 \in A_x^{ty_1+(1-t)y_2}$ and $(tY_1 + (1-t)Y_2) \cdot \xi(x) \geq \Phi_\xi(x, ty_1 + (1-t)y_2)$, which proves the convexity of $\Phi_\xi(x, \cdot)$.

3.) Let $\xi_n \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)^{\mathbb{N}}$ and $x \in \overline{\Omega}$ such that $\xi_n(x) \to \xi(x)$ as $n \to +\infty$. Still thanks to Assumption 2.4, there exists $Y_n \in A_x^y$ such that $|Y_n| \leq C|y|$ and $Y_n \cdot \xi_n(x) = \Phi\xi_n(x, y)$. Then up to a subsequence, there exists $Y \in A_x^y$ such that Y_n converges to Y. By definition, we have that $Y \cdot \boldsymbol{\xi}(x) \geq \Phi_\xi(x, y)$. Let $Y' \in A_x^y$ such that $Y' \cdot \boldsymbol{\xi}(x) = \Phi_\xi(x, y)$. Then we have $Y' \cdot \boldsymbol{\xi}_n(x) \geq Y_n \cdot \boldsymbol{\xi}_n(x)$ and passing to the limit, $Y' \cdot \boldsymbol{\xi}(x) \geq Y \cdot \boldsymbol{\xi}(x)$. So $\lim \Phi_{\xi_n}(x) = Y \cdot \xi(x) = \Phi_\xi(x)$. $\square$

Due to Lemma 2.1, we may reformulate (2.22) :

$$\widetilde{L}_\xi(\sigma) = \int_0^1 \Phi_\xi(\sigma(t), \dot{\sigma}(t))\, dt = \int_0^1 |\dot{\sigma}(t)| \Phi_\xi \left(\sigma(t), \frac{\dot{\sigma}(t)}{|\dot{\sigma}(t)|}\right) dt.$$

The following lemma gives a Hölder estimate for c_ξ.

Lemma 2.2. *There exists a nonnegative constant C such that for every $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ and every $(x_1, x_2, y_1, y_2) \in \Omega^4$, one has*

$$|c_\xi(x_1, y_1) - c_\xi(x_2, y_2)| \leq C\|\xi\|_{L^p(\theta)}(|x_1 - x_2|^\beta + |y_1 - y_2|^\beta), \tag{2.23}$$

where $\beta = 1 - d/p$. So if $(\xi_n)_n \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)^{\mathbb{N}}$ is bounded in $L^p(\theta)$, then $(c_{\xi_n})_n$ admits a subsequence that converges in $C(\overline{\Omega} \times \overline{\Omega}, \mathbb{R}_+)$.

Proof. Thanks to Morrey's inequality, it is sufficient to show

$$|\nabla_y c_\xi(x, \cdot)| \leq |\boldsymbol{\xi}(\cdot)|_1 \text{ for all } x \in \Omega$$

and

$$|\nabla_x c_\xi(\cdot, y)| \leq |\boldsymbol{\xi}(\cdot)|_1 \text{ for all } x \in \Omega,$$

where $|\boldsymbol{\xi}(x)|_1 = \sum_{k=1}^N \xi(x, v_k(x))$ for all $x \in \Omega$. Let $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, $(x, y) \in \Omega^2$. For every $k > 0$, let us define $\sigma_k \in C_{x,y}$ such that

$$\widetilde{L}_\xi(\sigma_k) \leq c_\xi(x, y) + \frac{1}{k}.$$

For all $\varepsilon > 0$ and let z such that $y + \varepsilon z \in \Omega$ and any $t_0 \in (0,1)$, define

$$\sigma_{k,t_0}(t) = \begin{cases} \sigma_k\left(\dfrac{t}{t_0}\right) & \text{for } t \in [0, t_0], \\ y + \left(\dfrac{t - t_0}{1 - t_0}\right)\varepsilon z & \text{for } t \in [t_0, 1]. \end{cases}$$

Then for all $k > 0$,

$$\begin{aligned}
c_\xi(x, y+\varepsilon z) &\le \int_0^1 \inf_{X\in A_{\sigma_{k,t_0}(t)}^{\dot\sigma_{k,t_0}(t)}} \left(\sum_{k=0}^N x_k \xi(\sigma_{k,t_0}(t), v_k(\sigma_{k,t_0}(t)))\right) dt\\
&= \int_0^{t_0} \inf_{X\in A_{\sigma_k(t/t_0)}^{1/t_0\dot\sigma_k(t/t_0)}} \left(\sum_{k=0}^N x_k \xi\left(\sigma_k\left(\frac{t}{t_0}\right), v_k\left(\frac{t}{t_0}\right)\right)\right) dt\\
&+ \int_{t_0}^1 \inf_{X\in A_{y+\varepsilon z(t-t_0)/(1-t_0)}^{\varepsilon z/(1-t_0)}} \left(\sum_{k=0}^N x_k \xi\left(y+\frac{t-t_0}{1-t_0}\varepsilon z, v_k\left(y+\frac{t-t_0}{1-t_0}\varepsilon z\right)\right)\right) dt\\
&= \widetilde{L}_\xi(\sigma_k) + \varepsilon \int_0^1 \inf_{X\in A_{y+t\varepsilon z}^z} \left(\sum_{k=1}^N x_k \xi(y+t\varepsilon z, v_k(y+t\varepsilon z))\right) dt
\end{aligned}$$

by making a change of variables for each term. Finally we have

$$c_\xi(x, y+\varepsilon z) \le c_\xi(x,y) + \frac{1}{k} + C\varepsilon|z| \int_0^1 |\boldsymbol{\xi}(y+t\varepsilon z, v)|\, dt$$

since for every $y \in \overline{\Omega}, z \in \mathbb{R}^d$, there exists $X \in A_y^z$ such that $|X| \le C|z|$ due to Assumption 2.4. Hence by passing to the limit, we obtain

$$|\nabla_y c_\xi(x,\cdot)| \le C|\boldsymbol{\xi}(\cdot)| \in L^p(\theta) \text{ for all } x \in \Omega.$$

The argument is the same for the other variable. So $c_\xi(x,\cdot)$ and $c_\xi(\cdot,y) \in W^{1,p}(\Omega)$. Since $p > d$, we deduce from Morrey's theorem and Assumption 2.3 that there is a constant $C > 0$ such that

$$\begin{aligned}
|c_\xi(x,y_1) - c_\xi(x,y_2)| &\le C\|\xi\|_{L^p(\theta)}|y_1-y_2|^\beta\ \forall x, y_1, y_2 \text{ in } \Omega,\\
|c_\xi(x_1,y) - c_\xi(x_2,y)| &\le C\|\xi\|_{L^p(\theta)}|x_1-x_2|^\beta\ \forall x_1, x_2, y \text{ in } \Omega.
\end{aligned}$$

This proves (2.23). We have the second statement in the lemma due to (2.23), the identity $c_{\xi_n}(x,x) = 0$ and Ascoli's theorem. □

For every $\xi \in L_+^p(\theta)$, let us define

$$\overline{c}_\xi(x,y) := \sup\{c(x,y) : c \in \mathcal{A}(\xi)\},\ \forall (x,y) \in \overline{\Omega}\times\overline{\Omega} \tag{2.24}$$

where

$$\mathcal{A}(\xi) := \{\lim_n c_{\xi_n} \text{ in } C(\overline{\Omega}\times\overline{\Omega}) : (\xi_n)_n \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R}_+)^{\mathbb{N}}, \xi_n \to \xi \text{ in } L^p(\theta)\}.$$

We will justify in the next section that for every $\xi \in L_+^p(\theta)$, the continuous limit functional is

$$J(\xi) := I_0(\xi) - I_1(\xi) := \int_{\Omega\times\mathbb{S}^{d-1}} H(x,v,\xi(x,v))\theta(dx,dv) - \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi d\gamma. \tag{2.25}$$

The following lemma is a generalization of Lemma 3.5 in [39]:

Lemma 2.3. *If $\xi \in L_+^p(\theta)$ then there exists a sequence $(\xi_n)_n$ in $C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R}_+)$ such that c_{ξ_n} converges to $\overline{c}_\xi$ in $C(\overline{\Omega}\times\overline{\Omega})$ as $n\to\infty$.*

The proof is the same as in [39]. We build a sequence $(\xi_n) \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ such that

$$\|\xi_n - \xi\|_{L^p(\theta)} \leq \frac{1}{n} \text{ and } |\overline{c}_\xi(x_k, y_k) - c_{\xi_n}(x_k, y_k)| \leq \frac{1}{n} \text{ for all } k \leq n,$$

where $(x_k, y_k)_{k\in\mathbb{N}}$ is a dense sequence of points in $\overline{\Omega} \times \overline{\Omega}$. Then we conclude with Lemma 2.2 and Ascoli's theorem.

When ξ is continuous, one has the following result that extends Lemma 3.4 in [39].

Lemma 2.4. *If $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ then $c_\xi = \overline{c}_\xi$.*

Proof. The inequality $\overline{c}_\xi \geq c_\xi$ is immediate since it is sufficient to take the constant sequence $\xi_n = \xi$ in the definition of $\overline{c}_\xi$. Conversely, let $x, y \in \Omega$, $\varepsilon > 0$ and $\sigma \in C_{x,y}$ such that $L_\xi(\sigma) < c_\xi(x, y) + 1/k$. We can choose σ piecewise linear by density of this kind of curves and using Lemma 2.1. Let $(S_i)_{i=0,\dots,m-1}$ be the segments which compose σ with $S_i = [x_i, x_{i+1}]$, $x_0 = x$ and $x_m = y$. Let a sequence $\xi_n \to \xi$ such that $c_{\xi_n} \to c$. We want to show $c \leq c_\xi$. Take a small $\delta > 0$. For every $\alpha \in [0, \delta]$ and any vector $\vec{u}$ unitary and perpendicular to $\overrightarrow{x_0 x_1}$, let us define a curve $\sigma_{\alpha,\vec{u}}$ in the following way. For every $i = 0, \dots, m-1$, let us consider the tube of radius δ

$$\begin{aligned} T_i(\delta) &= \{x + \vec{v}; x \in [x_i, x_{i+1}], < \vec{v}, \overrightarrow{x_i, x_{i+1}} >= 0 \text{ and } |\vec{v}| \leq \delta\} \\ &= \bigcup_{x\in[x_i,x_{i+1}]} D_i(x, \delta) = \{x + \vec{v}; < \vec{v}, \overrightarrow{x_i, x_{i+1}} >= 0 \text{ and } |\vec{v}| \leq \delta\}. \end{aligned}$$

Here $< \cdot, \cdot >$ is the dot product and $|\cdot|$ the Euclidean norm in $\mathbb{R}^d$. The disk $D_i(x, \delta)$ is a submanifold of dimension $d-1$. The intersection $H_{i+1}(\delta) = D_i(x_i, \delta) \cap D_{i+1}(x_{i+1}, \delta)$ is in an affine space of dimension $d-2$. Let us also define $\mathcal{V}_i$ the set of vectors that are unitary and perpendicular to $\overrightarrow{x_i x_{i+1}}$. First, let us define $x_0'(\alpha, \vec{u}) = x_0 + \alpha\vec{u}$ and $x_0''(\alpha, \vec{u}) = x_1 + \alpha\vec{u}$. Then $\sigma_{\alpha,\vec{u}}$ links $x_0'(\alpha, \vec{u})$ to $x_0''(\alpha, \vec{u})$ by the segment $S_0(\alpha, \vec{u}) = [x_0'(\alpha, \vec{u}), x_0''(\alpha, \vec{u})]$. Let $P_1(\alpha, \vec{u})$ denote the plane perpendicular to $H_1(\delta)$ and passing through $x_0''(\alpha, \vec{u})$. Then $P_1(\alpha, \vec{u}) \cap H_0(\delta)$ is a singleton denoted by $z_1(\alpha, \vec{u})$. We have that $\overrightarrow{z_1(\alpha, \vec{u}), x_0''(\alpha, \vec{u})}$ is perpendicular to $H_1(\delta)$. The set $P_1(\alpha, \vec{u}) \cap T_0(\delta)$, respectively $P_1(\alpha, \vec{u}) \cap T_1(\delta)$, is a segment passing through $z_1(\alpha, \vec{u})$ and contained in the line denoted by $\mathcal{D}_0'(\alpha, \vec{u})$, respectively $\mathcal{D}_1(\alpha, \vec{u})$. There exists an unique point denoted by $x_1'(\alpha, \vec{u})$ in $T_1(\delta)$ such that $|z_1(\alpha, \vec{u})x_0''(\alpha, \vec{u})| = |z_1(\alpha, \vec{u})x_1'(\alpha, \vec{u})|$ and $x_0''(\alpha, \vec{u})$ and $x_1'(\alpha, \vec{u})|$ are on the same side of $\mathcal{D}_1(\alpha, \vec{u})$ in $P_1(\alpha, \vec{u})$. Then $\sigma_{\alpha,\vec{u}}$ links $x_0''(\alpha, \vec{u})$ to $x_1'(\alpha, \vec{u})$ by the arc $A_1(\alpha, \vec{u})$ with center $z_1(\alpha, \vec{u})$ in this side of $P_1(\alpha, \vec{u})$. Let R_1 denote the isometry that maps $x_0''(\alpha, \vec{u})$ to $x_1'(\alpha, \vec{u})$ and that lets invariant $H_1(\delta)$. The vector $R_1(\vec{u})$ is in $\mathcal{V}_1$. We iterate the process on each of the segments $S_i, i = 1, \dots, m-1$. In this way we obtain $\sigma_{\alpha,\vec{u}} \in C^{x_{\alpha,\vec{u}},y_{\alpha,\vec{u}}}$ where $x_{\alpha,\vec{u}} = x_0'(\alpha, \vec{u})$ and $y_{\alpha,\vec{u}} = x_m''(\alpha, \vec{u})$. We define $t_i''(\alpha, \vec{u})$ and $t_{i+1}'(\alpha, \vec{u})$ to be the increasing sequences in $(0, 1)$ such that $\sigma_{\alpha,\vec{u}}([t_i''(\alpha, \vec{u}), t_{i+1}'(\alpha, \vec{u})]) = A_{i+1}(\alpha, \vec{u})$. We can take these sequences such that $t_i'(\alpha, \vec{u}) \to t_i'$ and $t_i''(\alpha, \vec{u}) \to t_i'' = t_{i+1}'$ as $\alpha \to 0^+$ for every $\vec{u} \in S_0$ where the t_i' are such that $\sigma(t_i') = x_i$.

Then we compute

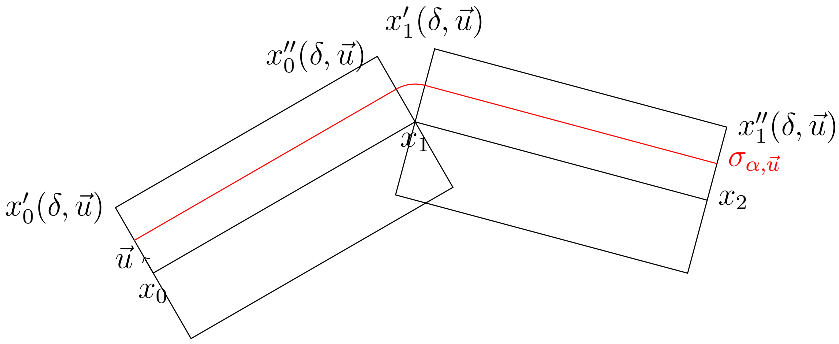

Figure 2.4 – An example for $d = 2$.

$$\int_0^\delta \left(\int_{\mathcal{V}_0} L_{\xi_n}(\sigma_{\alpha,\vec{u}})d\vec{u}\right) d\alpha = \sum_{i=0}^{m-1} \int_{T_i(\delta)} \Phi_{\xi_n}\left(\cdot, \frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right) d\mathcal{L}^d$$
$$+ \sum_{i=1}^{m-1} \int_0^\delta \left(\int_{\mathcal{V}_0} \left(\int_{t_i''(\alpha,\vec{u})}^{t_{i+1}'(\alpha,\vec{u})} \Phi_{\xi_n}(\sigma_{\alpha,\vec{u}}(t), \dot{\sigma}_{\alpha,\vec{u}}(t))dt\right) d\vec{u}\right) d\alpha.$$

Since $c_{\xi_n}(x_{\alpha,\vec{u}}, y_{\alpha,\vec{u}}) \leq L_{\xi_n}(\sigma_{\alpha,\vec{u}})$ we obtain

$$\int_0^\delta \left(\int_{\mathcal{V}_0} c_{\xi_n}(x_{\alpha,\vec{u}}, y_{\alpha,\vec{u}})d\vec{u}\right) d\alpha \leq \int_0^\delta \left(\int_{\mathcal{V}_0} L_{\xi_n}(\sigma_{\alpha,\vec{u}})d\vec{u}\right) d\alpha.$$

Moreover, up to a subsequence, we have that ξ_n converges to ξ in $L^p(\theta)$ and that $\boldsymbol{\xi_n}$ converges to $\boldsymbol{\xi}$ almost everywhere. Then due to the uniform convergence of c_{ξ_n} to c, Lemma 2.1 and the dominated convergence theorem, we have

$$\int_0^\delta \left(\int_{\mathcal{V}_0} c(x_{\alpha,\vec{u}}, y_{\alpha,\vec{u}})d\vec{u}\right) d\alpha \leq \sum_{i=0}^{m-1} \int_{T_i(\delta)} \Phi_{\xi}\left(\cdot, \frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right) d\mathcal{L}^d$$
$$+ \sum_{i=1}^{m-1} \int_0^\delta \left(\int_{\mathcal{V}_0} \left(\int_{t_i''(\alpha,\vec{u})}^{t_{i+1}'(\alpha,\vec{u})} \Phi_{\xi}(\sigma_{\alpha,\vec{u}}(t), \dot{\sigma}_{\alpha,\vec{u}}(t))dt\right) d\vec{u}\right) d\alpha.$$

Then we divide by δ and pass to the limit as $\delta \to 0^+$. Since c is continuous, we get

$$\lim_{\delta\to 0^+} \frac{1}{\delta} \int_0^\delta \left(\int_{\mathcal{V}_0} c(x_{\alpha,\vec{u}}, y_{\alpha,\vec{u}})d\vec{u}\right) d\alpha = \int_{\mathcal{V}_0} c(x_{0,\vec{u}}, y_{0,\vec{u}})d\vec{u} = \mathcal{A}_{d-1}c(x,y)$$

where $\mathcal{A}_{d-1}$ is the hypervolume of a $(d-2)$-dimensional unit sphere. On the other hand, in the second term, we integrate over a compact whose volum is $O(\delta^d)$ so that due to Assumption 2.4, Φ_ξ is bounded. Therefore we have for $\delta \to 0^+$

$$\frac{1}{\delta} \sum_{i=1}^{m-1} \int_0^\delta \left(\int_{\mathcal{V}_0} \left(\int_{t_i''(\alpha,\vec{u})}^{t_{i+1}'(\alpha,\vec{u})} \Phi_{\xi}(\sigma_{\alpha,\vec{u}}(t), \dot{\sigma}_{\alpha,\vec{u}}(t))dt\right) d\vec{u}\right) d\alpha$$
$$\leq mC\|\xi\|_\infty \delta^{d-1} \to 0.$$

For the first term in the right side, we have

$$
\begin{aligned}
&\lim_{\delta\to 0^+}\frac{1}{\delta}\int_0^\delta\sum_{i=0}^{m-1}\int_{T_i(\delta)}\Phi_\xi\left(\cdot,\frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right)d\mathcal{L}^d\\
&=\lim_{\delta\to 0^+}\sum_{i=0}^{m-1}\int_0^\delta\left(\int_{\mathcal{V}_0}\left(\int_{t_i'(\alpha,\vec{u})}^{t_i''(\alpha,\vec{u})}|\dot{\sigma}_{\alpha,\vec{u}}(t)|\Phi_\xi\left(\sigma_{\alpha,\vec{u}}(t),\frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right)dt\right)d\vec{u}\right)d\alpha\\
&=\sum_{i=0}^{m-1}\int_{\mathcal{V}_0}\left(\int_{t_i'(0,\vec{u})}^{t_i''(0,\vec{u})}|\dot{\sigma}_{0,\vec{u}}(t)|\Phi_\xi\left(\sigma_{0,\vec{u}}(t),\frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right)dt\right)d\vec{u}\\
&=\mathcal{A}_{d-1}\sum_{i=0}^{m-1}\int_{S_i}\Phi_\xi\left(\cdot,\frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}\right)d\mathcal{L}^1=\mathcal{A}_{d-1}L_\xi(\sigma).
\end{aligned}
$$

Finally we obtain

$$c(x,y)\le L_\xi(\sigma)<c_\xi(x,y)+\frac{1}{k}.$$

Since k is chosen arbitrarily, we get $c\le c_\xi$ and the desired result.

□

3.3 The Γ-convergence result

We will prove that the problem (2.25) is the continuous limit of the discrete problems (2.18) in the Γ-convergence sense. The Γ-convergence theory is a powerful tool to study the convergence of variational problems (convergence of values but also of minimizers) depending on a parameter. Here we want to study problems depending on a scale parameter (which is ε), it is particularly well suited. References for the general theory of Γ-convergence and many applications are the books of Dal Maso [80] and Braides [29].

First let us define weak L^p convergence of a discrete family $\xi^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}$.

Definition 2.2. *For $\varepsilon>0$, let $\xi^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}$ and $\xi\in L_+^p(\theta)$, then ξ^ε is said to weakly converge to ξ in L^p ($\xi^\varepsilon\rightharpoonup\xi$) if :*

1. *There exists a constant $M>0$ such that for all $\varepsilon>0$, one has*

$$\|\xi^\varepsilon\|_{\varepsilon,p}:=\left(\sum_{(x,e)\in E^\varepsilon}|e|^d\xi^\varepsilon(x,e)^p\right)^{1/p}\le M.$$

2. *For every $\varphi\in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R})$, one has*

$$\lim_{\varepsilon\to 0^+}\sum_{(x,e)\in E^\varepsilon}|e|^d\varphi\left(x,\frac{e}{|e|}\right)\xi^\varepsilon(x,e)=\int_{\Omega\times\mathbb{S}^{d-1}}\varphi(x,v)\xi(x,v)\theta(dx,dv).$$

Definition 2.3. *For $\varepsilon>0$, let $F^\varepsilon:\mathbb{R}_+^{\#E^\varepsilon}\to\mathbb{R}\cup\{+\infty\}$ and $F:L_+^p(\theta)\to\mathbb{R}\cup\{+\infty\}$, then the family of functionals $(F^\varepsilon)_\varepsilon$ is said to Γ-converge (for the weak L^p topology) to F if and only if the following two conditions are satisfied:*

1. *(Γ-liminf inequality) For every $\xi \in L^p_+(\theta)$ and $\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}$ such that $\xi^\varepsilon \to \xi$, one has*
$$\liminf_{\varepsilon\to 0^+} F^\varepsilon(\xi^\varepsilon) \geq F(\xi),$$
2. *(Γ-limsup inequality) For every $\xi \in L^p_+(\theta)$, there exists $\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}$ such that $\xi^\varepsilon \to \xi$ and*
$$\limsup_{\varepsilon\to 0^+} F^\varepsilon(\xi^\varepsilon) \leq F(\xi),$$

Now, we can state our main result, whose complete proof will be performed in the next section:

Theorem 2.2. *Under all previous assumptions, the family of functionals $(J^\varepsilon)_\varepsilon$ Γ-converges (for the weak L^p topology) to the functional J defined by* (2.25)*.*

Classical arguments from general Γ-convergence theory allow us to have the following convergence result:

Corollary 2.1. *Under all previous assumptions, the problems* (2.18) *for all $\varepsilon > 0$ and* (2.25) *admit solutions and one has:*
$$\min_{\xi^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}} J^\varepsilon(\xi^\varepsilon) \to \min_{\xi\in L^p_+(\theta)} J(\xi).$$
Moreover, if for any $\varepsilon > 0$, ξ^ε is the solution of the minimization problem (2.18) *then $\xi^\varepsilon \to \xi$ where ξ is the minimizer of J over $L^p_+(\theta)$.*

Proof. First, due to (2.13) we have
$$\begin{aligned}I_0^\varepsilon(\xi^\varepsilon) &\geq \lambda \sum_{(x,e)\in E^\varepsilon} |e|^d \left(\xi^\varepsilon(x,e)^p - 1\right)\\ &= \lambda\|\xi^\varepsilon\|^p_{\varepsilon,p} - \lambda\sum |e|^d\\ &\geq \lambda\|\xi^\varepsilon\|^p_{\varepsilon,p} - C,\end{aligned}$$
since from Assumption 2.5 it follows that
$$\sum_{(x,e)\in E^\varepsilon} |e|^d \to \int_{\Omega\times\mathbb{S}^{d-1}} \theta(dx,dv) = \sum_{k=1}^N \int_\Omega c_i(x)\,dx = C.$$
To estimate the other term $I_1^\varepsilon(\xi^\varepsilon)$, let us write
$$c^\varepsilon(x,y) := \min_{\sigma\in C^\varepsilon_{x,y}} \sum_{(z,e)\subset\sigma} |e|\xi^\varepsilon(z,e).$$
$\forall x_0 \in N^\varepsilon, \forall x,y \in N^\varepsilon$ two neighboring nodes, we have
$$\begin{aligned}|c^\varepsilon(x_0,x) - c^\varepsilon(x_0,y)| &\leq \max_{e/(x,e)\in E^\varepsilon} |e|\xi^\varepsilon(x,e)\\ &\leq \varepsilon \max_{e/(x,e)\in E^\varepsilon} \xi^\varepsilon(x,e).\end{aligned}$$
Then thanks to Lemma 2.6, we have for every $x,y \in N^\varepsilon$,
$$|c^\varepsilon(x_0,x) - c^\varepsilon(x_0,y)| \leq C\left\|\max_{e/(\cdot,e)\in E^\varepsilon} \xi^\varepsilon(\cdot,e)\right\|_{\varepsilon,p},$$

hence with $x_0 = x$, thanks to Assumption 2.1, we obtain

$$\begin{aligned} c^\varepsilon(x,y) &\le C \left\| \max_{e/(\cdot,e)\in E^\varepsilon} \xi^\varepsilon(\cdot,e) \right\|_{\varepsilon,p} \\ &\le C\|\xi^\varepsilon\|_{\varepsilon,p} \end{aligned}$$

so that recalling (2.20), we have

$$|I_1^\varepsilon(\xi^\varepsilon)| \le C\|\xi^\varepsilon\|_{\varepsilon,p} \sum_{x,y\in N^\varepsilon} \varepsilon^{d/2-1}\gamma^\varepsilon(x,y) \le C\|\xi^\varepsilon\|_{\varepsilon,p},$$

because it follows from Assumption 2.8 that

$$\sum_{x,y\in N^\varepsilon} \varepsilon^{d/2-1}\gamma^\varepsilon(x,y) \to \int_{\Omega\times\Omega} d\gamma \text{ as } \varepsilon \to 0^+.$$

In particular, we get the equi-coercivity estimate

$$J^\varepsilon(\xi^\varepsilon) \ge C(\|\xi^\varepsilon\|_{\varepsilon,p}^p - \|\xi^\varepsilon\|_{\varepsilon,p} - 1).$$

Since J^ε is continuous on $\mathbb{R}_+^{\#E^\varepsilon}$, this proves that the infimum of J^ε over $\mathbb{R}_+^{\#E^\varepsilon}$ is attained at some $\widetilde{\xi^\varepsilon}$ and also that $\|\widetilde{\xi^\varepsilon}\|_{\varepsilon,p}$ is bounded. Moreover, H is strictly convex thanks to Assumption 2.9 and I_0^ε is concave since it is defined through minimizations of concave functions. So J^ε is strictly convex and therefore $\widetilde{\xi^\varepsilon}$ is unique. In particular, we can define for $\varepsilon > 0$ the following Radon measure M_ε :

$$\langle M_\varepsilon, \varphi\rangle := \sum_{(x,e)\in E^\varepsilon} |e|^d \varphi\left(x, \frac{e}{|e|}\right) \widetilde{\xi^\varepsilon}(x,e),\ \forall\varphi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R}).$$

Due to Hölder inequality we have for every $\varepsilon > 0$ and $\varphi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R})$,

$$|\langle M_\varepsilon, \varphi\rangle| \le C\|\varphi\|_{\varepsilon,q} \tag{2.26}$$

where $q = p/(p-1)$ is the conjugate exponent of p and the semi-norm $\|.\|_{\varepsilon,q}$ is defined by:

$$\|\varphi\|_{\varepsilon,q} := \left(\sum_{(x,e)\in E^\varepsilon} |e|^d \left|\varphi\left(x, \frac{e}{|e|}\right)\right|^q \right)^{1/q}.$$

Because of Assumption 2.3 there is a nonnegative constant C such that $\|\varphi\|_{\varepsilon,q} \le C\|\varphi\|_\infty$ for every $\varphi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R})$. We deduce from (2.26) and Banach-Alaoglu's theorem that there exists a subsequence (still denoted by M_ε) and a Radon measure M over $\overline{\Omega}\times\mathbb{S}^{d-1}$ with values in $\mathbb{R}$ to which M_ε weakly star converges. Moreover, for every $\varphi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R})$, due to (2.26) and Assumption 2.3 (with $|\varphi|^q$), we have

$$|\langle M, \varphi\rangle| \le C \lim_{\varepsilon\to 0^+} \|\varphi\|_{\varepsilon,q} = C\|\varphi\|_{L^q(\theta)}$$

which proves that M in fact admits an Ł$^p p(\theta)$-representative denoted by $\widetilde{\xi}$. Besides $\widetilde{\xi} \in L_+^p(\theta)$ (componentwise nonnegativity is stable under weak convergence) and

$\widetilde{\xi}^\varepsilon \to \widetilde{\xi}$ in the sense of definition 2.2. We still have to prove that $\widetilde{\xi}$ minimizes J over $L^p_+(\theta)$. First from the Γ-liminf inequality we find that

$$J(\widetilde{\xi}) \leq \liminf_{\varepsilon \to 0^+} J^\varepsilon(\xi^\varepsilon) = \liminf_{\varepsilon \to 0^+} \min_{\xi^\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}} J^\varepsilon(\xi^\varepsilon).$$

Let $\zeta \in L^p_+(\theta)$, we know from the Γ-limsup inequality that there exists a sequence $(\zeta^\varepsilon)_\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}$ such that $\zeta^\varepsilon \to \zeta$ in the sense of definition 2.2 and that

$$\limsup_{\varepsilon \to 0^+} J^\varepsilon(\zeta^\varepsilon) \leq J(\zeta).$$

Since $\widetilde{\xi}^\varepsilon$ minimizes J^ε we have that

$$J(\widetilde{\xi}) \leq \liminf_{\varepsilon \to 0^+} J^\varepsilon(\xi^\varepsilon) \leq \limsup_{\varepsilon \to 0^+} J^\varepsilon(\xi^\varepsilon) \leq \limsup_{\varepsilon \to 0^+} J^\varepsilon(\zeta^\varepsilon) \leq J(\zeta).$$

We can then deduce that $\widetilde{\xi}$ minimizes J over $L^p_+(\theta)$ and we have also proved the existence of a minimizer to the limit problem. We also have that

$$\min_{L^p_+(\theta)} J \leq \liminf_{\varepsilon \to 0^+} \min_{\xi^\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}} J^\varepsilon(\xi^\varepsilon) \leq \limsup_{\varepsilon \to 0^+} \min_{\xi^\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}} J^\varepsilon(\xi^\varepsilon) \leq J(\zeta), \ \forall \zeta \in L^p_+(\theta)$$

which provides the convergence of the values of the discrete minimization problems to the value of the continuous one. Furthermore we have convergence of the whole family $\widetilde{\xi}^\varepsilon$ and not only of a subsequence by the uniqueness of the minimizer $\widetilde{\xi}$ of J over $L^p_+(\theta)$ since J is strictly convex (by the same reasoning as for J^ε). $\square$

4 Proof of the theorem

4.1 The Γ-liminf inequality

For $\varepsilon > 0$, let $\xi^\varepsilon \in \mathbb{R}_+^{\# E^\varepsilon}$ and $\xi \in L^p_+(\theta)$ such that $\xi^\varepsilon \to \xi$ (in the sense of definition 2.2). In this subsection, we want to prove that

$$\liminf_{\varepsilon \to 0^+} J^\varepsilon(\xi^\varepsilon) \geq J(\xi). \tag{2.27}$$

We need some lemmas to establish this inequality. The first one concerns the terms I_0^ε and I_0.

Lemma 2.5. *One has*

$$\liminf_{\varepsilon \to 0^+} I_0^\varepsilon(\xi^\varepsilon) \geq I_0(\xi).$$

Proof. Let $\delta > 0$, then there exists φ continuous on $\overline{\Omega} \times \mathbb{S}^{d-1}$ such that

$$I_0(\xi) \leq \delta + \int_{\Omega \times \mathbb{S}^{d-1}} (\varphi(x,v)\xi(x,v) - G(x,v,\varphi(x,v)))\theta(dx,dv),$$

where $G(x,v,\cdot)$ is the Legendre transform of $H(x,v,\cdot)$. Indeed, we have the existence of such a function (not necessarily continuous) by a convex duality argument and we have the continuity thanks to the continuity for the L^q-topology of

$\varphi \to \int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,\varphi(x,v))\theta(dx,dv)$ (it follows from Krasnoselskii's theorem [72]) and the density of continuous functions in L^q.

Now from Young's inequality we obtain for every $\varepsilon > 0$ and $(x,e) \in E^\varepsilon$

$$H\left(x, \frac{e}{|e|}, \xi^\varepsilon(x,e)\right) \geq \varphi\left(x, \frac{e}{|e|}\right)\xi^\varepsilon(x,e) - G\left(x, \frac{e}{|e|}, \varphi\left(x, \frac{e}{|e|}\right)\right).$$

We deduce that

$$\liminf_{\varepsilon\to 0^+} I_0^\varepsilon(\xi^\varepsilon) \geq I_0(\xi) - \delta.$$

Since $\delta > 0$ is chosen arbitrarily, we obtain the desired result. $\square$

Now we need the following discrete version of Morrey's inequality to have information about the nonlocal term.

Lemma 2.6. *Let $\theta^\varepsilon \in \mathbb{R}_+^{\#N^\varepsilon}$ and $\varphi^\varepsilon \in \mathbb{R}_+^{\#N^\varepsilon}$ such that*

$$|\varphi^\varepsilon(x) - \varphi^\varepsilon(y)| \leq \varepsilon\theta^\varepsilon(x), \text{ for every } x \in N^\varepsilon \text{ and every } y \text{ neighbor of } x, \tag{2.28}$$

then there exists a constant C such that for every $(x,y) \in E^\varepsilon \times E^\varepsilon$, one has

$$|\varphi^\varepsilon(x) - \varphi^\varepsilon(y)| \leq C\|\theta^\varepsilon\|_{\varepsilon,p}|x-y|^\beta$$

where $\beta = 1 - d/p$ and

$$\|\theta^\varepsilon\|_{\varepsilon,p} := \left(\varepsilon^d \sum_{x\in N^\varepsilon} \theta^\varepsilon(x)^p\right)^{1/p}.$$

Proof. The idea is to linearly interpolate φ^ε in order to have a function in $W^{1,p}(\Omega)$ and then to apply Morrey's inequality. By recalling Assumption 2.2, let $V^\varepsilon = \mathrm{Conv}(x_1^\varepsilon, ..., x_L^\varepsilon)$ a polytope in the discrete network Ω_ε where for $k = 1, \dots, L$, x_k^ε is an neighbor of x_{k+1}^ε in N^ε, the indices being taken modulo L. Let us denote X^ε the isobarycenter of all these nodes (it is in the interior of V^ε by assumption) and let us define

$$\varphi^\varepsilon(X^\varepsilon) := \sum_{k=1}^{L} \frac{\varphi^\varepsilon(x_k^\varepsilon)}{L}.$$

Let $\{F_j^\varepsilon\}$ denote the subpolytopes given by Assumption 2.2 for V^ε with $F_j^\varepsilon = \mathrm{Conv}(X^\varepsilon, X_{j_1}^\varepsilon, \dots, X_{j_d}^\varepsilon)$ and these $(d+1)$ points that are linearly independent. Then for every x in V^ε, there exists j such that $x \in F_j^\varepsilon$. x is a conical combination of $X^\varepsilon, X_{j_1}^\varepsilon, \dots, X_{j_d}^\varepsilon$. There exists some unique nonnegative coefficients $\lambda, \lambda_1, \dots, \lambda_d \geq 0$ such that $x = \lambda X^\varepsilon + \sum_{i=1}^d \lambda_i X_{j_d}^\varepsilon$. We set

$$\varphi^\varepsilon(x) = \lambda\varphi(X^\varepsilon) + \sum_{i=1}^{d} \lambda_i \varphi\left(X_{j_d}^\varepsilon\right). \tag{2.29}$$

We still denote this interpolation by φ^ε. We then have $\varphi^\varepsilon \in W^{1,p}(\Omega)$ with $\|\nabla\varphi^\varepsilon\|_p \leq C\|\theta^\varepsilon\|_{\varepsilon,p}$ (computing $\nabla\varphi^\varepsilon$ is technical and we detail it below). We conclude thanks to Morrey's inequality.

Computing $\nabla\varphi^\varepsilon$:

We take the notations used in the above proof but we remove the ε-dependence for the sake of simplicity. Taking into account the construction of φ^ε, we will compute $\nabla\varphi^\varepsilon$ on a subpolytope $F = \mathrm{Conv}(X, X_1, ..., X_d)$ of V where X is the isobarycenter of all nodes x_k in V^ε and $X, X_1, \dots, X_d$ are linearly independent. We rearrange the x_k's such that x_k is a neighbor of x_{k+1}. Let $\mathcal{H}$ be the affine hyperplane

$$\mathcal{H} = \begin{pmatrix} X \\ \varphi^\varepsilon(X) \end{pmatrix} + \left\langle \begin{pmatrix} X_i - X \\ \varphi^\varepsilon(X_i) - \varphi^\varepsilon(X) \end{pmatrix} : i = 1, \dots, d \right\rangle.$$

The set $\mathcal{H}$ is the affine subspace generated by the graph of φ^ε on the subpolytope F. It is a hyperplane since the points $X, X_1, \dots, X_d$ are linearly independent. Then there exists some constants $a_0, \dots, a_{d+1}$ such that $\mathcal{H} = \{(z_1, \dots, z_{d+1}) \in \mathbb{R}^{d+1}; a_1 z_1 + \cdots + a_{d+1} z_{d+1} + a_0 = 0\}$ with $a_{d+1} \neq 0$ (otherwise $X, X_1, \dots, X_d$ would not be independent). The normal vector to the hyperplane is $\vec{n} = (a_1, \dots, a_{d+1})$ and

$$\nabla\varphi^\varepsilon = \begin{pmatrix} -a_1/a_{d+1} \\ \vdots \\ -a_d/a_{d+1} \end{pmatrix}.$$

Without loss of generality we assume $X = 0, X_1 = (y_1, 0), X_2 = (y_1^2, y_2, 0), \dots, X_d = (y_1^d, \dots, y_{d-1}^d, y_d)$, with $y_1, \dots, y_d \neq 0$. Then we have for $k = 1, \dots, d$

$$\begin{pmatrix} X_k - X \\ \varphi^\varepsilon(X_k) - \varphi^\varepsilon(X) \end{pmatrix} \cdot \vec{n} = 0 = \sum_{i<k} a_i y_i^k + a_k y_k + a_{d+1}(\varphi^\varepsilon(X_k) - \varphi^\varepsilon(X)),$$

i.e.

$$-\frac{a_k}{a_{d+1}} = \sum_{i<k} \frac{a_i}{a_{d+1}} \frac{y_i^k}{y_k} + \frac{\varphi^\varepsilon(X_k) - \varphi^\varepsilon(X)}{y_k}. \tag{2.30}$$

We must now find an estimate on each of these terms. First for $k = 1, \dots, d$, there exists $i_k \in \{1, \dots, L\}$ such that $X_k = x_{i_k}$ and it follows from (2.28) that we have

$$\begin{aligned} |\varphi^\varepsilon(X_k) - \varphi^\varepsilon(X)| &= \left| \sum_{l=1}^{L} \frac{L - l + 1}{L} (\varphi^\varepsilon(x_{l+i_k-1})) - \varphi^\varepsilon(x_{l+i_k})) \right| \\ &\leq \varepsilon \sum_{l=1}^{L} \theta^\varepsilon(x_l) \end{aligned}$$

with the indices taken modulo L. Moreover, due to (2.7), for $k = 1, \dots, d$ and $i < k$, we have $|y_i^k| \leq |X - X_k| \leq C\varepsilon$. Taking into account the fact that the X_k's are linearly independent, we get $|y_k| \geq C\varepsilon$. Then from (2.30) it follows for $k = 1, \dots, d$

$$\left| \frac{a_k}{a_{d+1}} \right| \leq C \sum_{i<k} \left| \frac{a_i}{a_{d+1}} \right| + C \sum_{l=1}^{L} \theta^\varepsilon(x_l). \tag{2.31}$$

We finally obtain by an induction on k that

$$\left| \frac{a_k}{a_{d+1}} \right| \leq C \sum_{l=1}^{L} \theta^\varepsilon(x_l)$$

so that due to (2.29), $|\nabla\varphi^\varepsilon(x)| \leq C \sum_{l=1}^{L} \theta^\varepsilon(x_l)$ for every x in the subpolytope F. We finally conclude with Assumption 2.2 that

$$\|\nabla\varphi^\varepsilon\|_p \leq C \|\theta^\varepsilon\|_{\varepsilon,p},$$

which completes the proof. □

The discretization of the Morrey inequality is crucial since we may now extend c^ε on $\Omega \times \Omega$. For every $(x, y) \in N^\varepsilon \times N^\varepsilon$, we define

$$c^\varepsilon(x, y) := \min_{\sigma \in C^\varepsilon_{x,y}} \sum_{(z,e)\subset\sigma} |e|\xi^\varepsilon(z, e) = \inf_{\sigma \in C^\varepsilon_{x,y}} \int_0^1 \Psi^\varepsilon(\tilde{\sigma}(t), \dot{\tilde{\sigma}}(t)) \cdot \boldsymbol{\xi}^\varepsilon(\tilde{\sigma}(t))dt. \tag{2.32}$$

By definition, if $x_0 \in \Omega_\varepsilon$ and x and y neighbors in Ω_ε, we have

$$c^\varepsilon(x_0, x) \leq c^\varepsilon(x_0, y) + \varepsilon \max_{e/(y,e)\in E^\varepsilon} \xi^\varepsilon(y, e).$$

Since $\|\xi^\varepsilon\|_{\varepsilon,p}$ is bounded, we deduce from Lemma 2.6 that there exists a constant C such that for every $\varepsilon > 0$ we have

$$|c^\varepsilon(x, y) - c^\varepsilon(x_0, y_0)| \leq C(|x - x_0|^\beta + |y - y_0|^\beta), \ \forall\, (x, y, x_0, y_0) \in (N^\varepsilon)^4.$$

We can then extend c^ε to the whole $\overline{\Omega} \times \overline{\Omega}$ (we still denote by c^ε this extension) by

$$c^\varepsilon(x, y) := \sup_{(x_0,y_0)\in\Omega_\varepsilon\times\Omega_\varepsilon} \{c^\varepsilon(x_0, y_0) - C(|x - x_0|^\beta + |y - y_0|^\beta)\}, \ \forall\, (x, y) \in \overline{\Omega} \times \overline{\Omega}.$$

By construction, c^ε still satisfy the uniform Hölder estimate on the whole $\overline{\Omega}\times\overline{\Omega}$ and since c^ε vanishes on the diagonal of $\overline{\Omega} \times \overline{\Omega}$, it follows from Arzela-Ascoli theorem that the family $(c^\varepsilon)_\varepsilon$ is relatively compact in $C(\overline{\Omega} \times \overline{\Omega})$. Up to a subsequence, we may therefore assume that there is some $c \in C(\overline{\Omega} \times \overline{\Omega})$ such that

$$c^\varepsilon \to c \text{ in } C(\overline{\Omega} \times \overline{\Omega}) \text{ and } c(x, x) = 0 \text{ for every } x \in \overline{\Omega}. \tag{2.33}$$

From Assumption 2.8 it may be concluded that

$$I_1^\varepsilon(\xi^\varepsilon) \leq \sum_{(x,y)\in N^\varepsilon\times N^\varepsilon} \varepsilon^{\frac{d}{2}-1} c^\varepsilon(x, y)\gamma^\varepsilon(x, y) \to \int_{\overline{\Omega}\times\overline{\Omega}} cd\gamma.$$

In consequence, with Lemma 2.5, it remains to prove $c \leq \overline{c}_\xi$ on $\overline{\Omega} \times \overline{\Omega}$. We will show that c is a sort of subsolution in a highly weak sense of an Hamilton-Jacobi equation and we will then conclude by some comparison principle. The end of this paragraph provides a proof of this inequality.

Lemma 2.7. *Let $x_0 \in \Omega$, $\xi \in L^p_+(\theta)$ and $\varphi \in W^{1,p}(\Omega)$ such that $\varphi(x_0) = 0$ (which makes sense since $p > d$ so that φ is continuous). If for a.e. $x \in \Omega$ one has*

$$\nabla\varphi(x) \cdot u \leq \Phi_\xi(x, u) := \inf_{U=(u_1,\dots,u_N)\in A^u_x} \left(\sum_{k=1}^N u_k \xi(x, v_k(x))\right) \text{ for all } u \in \mathbb{R}^d \tag{2.34}$$

then $\varphi \leq \overline{c}_\xi(x_0, \cdot)$ on Ω.

Remark 2.2. *The above assumption* (2.34) *is equivalent to :*

$$\nabla\varphi(x) \cdot v_k(x) \leq \xi_k(x) := \xi(x, v_k(x)), \ \forall x \text{ a.e.}, \forall k = 1, \dots, N.$$

Proof. The result is immediate if $\varphi \in C^1(\overline{\Omega})$ and ξ is continuous on $\overline{\Omega}$. Indeed, in this case, assumption (2.34) is true pointwise and if $x \in \Omega$ and σ is an absolutely continuous curve with values in $\overline{\Omega}$ connecting x_0 and x then by the chain rule we obtain

$$\varphi(x) = \int_0^1 \nabla\varphi(\sigma(t)) \cdot \dot{\sigma}(t)dt \leq \int_0^1 \Phi_\varepsilon(\sigma(t), \dot{\sigma}(t))\, dt$$

and taking the infimum in σ we get $\varphi \leq c_\xi(x_0, \cdot)$ on Ω so that $\varphi \leq \bar{c}_\xi(x_0, \cdot)$ due to Lemma 2.4. For the general case, if φ is only $W^{1,p}(\Omega)$ and ξ only $L^p_+(\theta)$, we first extend φ to a function in $W^{1,p}(\mathbb{R}^d)$ and we extend ξ outside Ω by writing $\xi(x, v) = |\nabla\varphi|(x)$ for every $x \in \mathbb{R}^d, v \in \mathbb{S}^{d-1}$ so that if $x \in \mathbb{R}^d \backslash \overline{\Omega}$ and $u \in \mathbb{S}^{d-1}$ we have

$$\begin{aligned}\nabla\varphi(x) \cdot u &\leq |\nabla\varphi(x)| \\ &\leq |\nabla\varphi(x)| \|\widetilde{U}\|_1 = \Phi_\xi(x, u),\end{aligned}$$

where $\widetilde{U} = (\tilde{u}_1, \ldots, \tilde{u}_N) \in A^u_x$ is a minimizer of $\Phi_\xi(x, u)$ The fact $U \in A^u_x$ implies that $|u| = 1 \leq \|U\|_1$. By homogeneity of (2.34) in u, (2.34) thus continues to hold outside Ω with the previous extensions. We then regularize φ and ξ. Let us take a mollifying sequence $\rho_n(x) = n^d\rho(nx)$, $x \in \mathbb{R}^d$ where ρ is a smooth nonnegative function supported on the unit ball and such that $\int_{\mathbb{R}^d} \rho = 1$. Set $\xi^n := \rho_n \star \xi$ and $\varphi_n := \rho_n \star \varphi - (\rho_n \star \varphi)(x_0)$. Let $x \in \mathbb{R}^d$. Recalling Remark 2.2 and the fact that the v_k's are α-Hölder continuous (Assumption 2.3), we have

$$\begin{aligned}\nabla\varphi_n(x) \cdot v_k(x) &= \int_{\mathbb{R}^d} \rho_n(y)\nabla\varphi(x-y) \cdot v_k(x)\, dy \\ &\leq \int_{\mathbb{R}^d} \rho_n(y)\xi_k(x-y)\, dy \\ &+ \int_{\mathbb{R}^d} \rho_n(y)\nabla\varphi(x-y) \cdot (v_k(x) - v_k(x-y))\, dy \\ &\leq \xi^n_k(x) + n^{d-\alpha}\|\rho\|_\infty \int_{B(0,1/n)} |\nabla\varphi(x-y)|\, dy \\ &\leq \xi^n_k(x) + Cn^{d-\alpha-d/q}\|\rho\|_\infty \|\nabla\varphi\|_p \\ &= \xi^n_k(x) + \varepsilon_n,\end{aligned}$$

where $\xi_k(x) = \xi(x, v_k(x))$ and $\varepsilon_n > 0$ such that $\varepsilon_n \to 0$ as $n \to \infty$ (since $\alpha > d/p$). So by using the above remark and the previous case where φ and ξ were regular, we have $\varphi_n \leq c_{\xi^n+\varepsilon_n}(x_0, .)$ and from the convergence of φ_n to φ it follows that

$$\varphi = \limsup \varphi_n \leq \limsup c_{\xi^n+\varepsilon_n}(x_0, \cdot) \leq \bar{c}_\xi(x_0, \cdot),$$

where the last inequality is given by the definition of $\bar{c}_\xi$ as a supremum (2.24) and the relative compactness of $c_{\xi^n+\varepsilon_n}$ in $C(\overline{\Omega} \times \overline{\Omega})$. □

We want to apply Lemma 2.7 to $c(x_0, \cdot)$ so that we need $c(x_0, \cdot) \in W^{1,p}(\Omega)$ for every $x_0 \in \Omega$. Let $(e_1, \ldots, e_d)$ given by Assumption 2.6, $\varphi \in C^1_c(\Omega)$ and $x^\varepsilon_0 \in \Omega_\varepsilon$ such that $|x_0 - x^\varepsilon_0| \leq \varepsilon$. Using the uniform convergence of $c^\varepsilon(x^\varepsilon_0, \cdot)$ to $c(x_0, \cdot)$ and Assumption 2.5, for $\varphi \in C^1_c(\Omega)$ and $i = 1, \ldots, d$ we have

$$\begin{aligned}T_i\varphi &:= \int_\Omega c(x_0, x)\nabla\varphi(x) \cdot e_i(x)\, dx \\ &= \lim_{\varepsilon\to 0^+} \sum_{\sigma \in C^\varepsilon_i} \sum_{k=0}^{N(\sigma)-1} |y_{k+1} - y_k|^d c^\varepsilon(x^\varepsilon_0, y_k) \frac{\varphi(y_{k+1}) - \varphi(y_k)}{|y_{k+1} - y_k|}\end{aligned}$$

where $\sigma = (y_0, \dots, y_{N(\sigma)})$. Then we can rearrange the sums as follows

$$\begin{aligned}
&T_i\varphi = \\
&\quad \lim_{\varepsilon\to 0^+} \sum_{\sigma\in C_i^\varepsilon} \left(\sum_{k=1}^{N(\sigma)-1} \varphi(y_k) \left(|y_k - y_{k-1}|^{d-1} c^\varepsilon(x_0^\varepsilon, y_{k-1}) - |y_{k+1} - y_k|^{d-1} c^\varepsilon(x_0^\varepsilon, y_k) \right) \right. \\
&\quad \left. + \varphi\left(y_{N(\sigma)}\right) |y_{N(\sigma)} - y_{N(\sigma)-1}|^{d-1} c^\varepsilon(x_0^\varepsilon, y_{N(\sigma)-1}) - \varphi(y_0)|y_1 - y_0|^{d-1} c^\varepsilon(x_0^\varepsilon, y_0) \right) \\
&\quad = \lim_{\varepsilon\to 0^+} \sum_{\sigma\in C_i^\varepsilon} \sum_{k=1}^{N(\sigma)-1} \varphi(y_k) \left(|y_k - y_{k-1}|^{d-1} c^\varepsilon(x_0^\varepsilon, y_{k-1}) - |y_{k+1} - y_k|^{d-1} c^\varepsilon(x_0^\varepsilon, y_k) \right)
\end{aligned}$$

since for ε small enough, y_0 and $y_{N(\sigma)}$ are not in the support of φ thanks to Assumption 2.6. For $\sigma \in C_i^\varepsilon$, we thus have

$$\begin{aligned}
&\sum_{k=1}^{N(\sigma)-1} \varphi(y_k) \left(|y_k - y_{k-1}|^{d-1} c^\varepsilon(x_0^\varepsilon, y_{k-1}) - |y_{k+1} - y_k|^{d-1} c^\varepsilon(x_0^\varepsilon, y_k) \right) \\
&\qquad = \sum_{k=1}^{N(\sigma)-1} \Big(\varphi(y_k)[|y_k - y_{k-1}|^{d-1}(c^\varepsilon(x_0^\varepsilon, y_{k-1}) - c^\varepsilon(x_0^\varepsilon, y_k)) \\
&\qquad + c^\varepsilon(x_0^\varepsilon, y_k)(|y_k - y_{k-1}|^{d-1} - |y_{k+1} - y_k|^{d-1})] \Big)
\end{aligned}$$

In the first term, we use the fact that if x and y are neighbors in Ω_ε then

$$c^\varepsilon(x_0^\varepsilon, x) \leq c^\varepsilon(x_0^\varepsilon, y) + |x - y|\xi^\varepsilon(y, x - y)$$

and we make an approximation on φ. Therefore this term is less than

$$\sum_{k=1}^{N(\sigma)-1} (1 + C|y_k - y_{k-1}|)|y_k - y_{k-1}|^d |\varphi(y_{k-1})| \xi^\varepsilon(y_{k-1}, y_k - y_{k-1}).$$

Due to Assumption 2.6 and the fact that the functions c^ε converge in $C(\overline{\Omega} \times \overline{\Omega})$ we obtain the following upper bound on the second term:

$$M(\varepsilon) \sum_{k=1}^{N(\sigma)-1} |y_k - y_{k-1}|^d |\varphi(y_k)|$$

where $M(\varepsilon) \geq 0$ and $\to 0$ as $\varepsilon \to 0$. Hence by using Hölder, the fact that $\|\xi^\varepsilon\|_{\varepsilon,p}$ is bounded and Assumption 2.3, we obtain

$$\left| \int_\Omega c(x_0, \cdot)\nabla\varphi \cdot e_i \right| \leq C\|\varphi\|_{Lq},\ \forall \varphi \in C_c^1(\Omega).$$

Since the functions $e_1, \dots, e_d$'s are continuous and linearly independent in $\mathbb{R}^d$ (see Assumption 2.6), this proves that $c(x_0, \cdot) \in W^{1,p}(\Omega)$. By a similar argument we obtain that $c(\cdot, y_0) \in W^{1,p}(\Omega)$ for every $y_0 \in \Omega$.

Remark 2.3. *Recalling Assumption 2.6 let the measures* $\theta_i \in \mathcal{M}_+(\Omega \times \mathbb{S}^{d-1})$, $i = 1, \dots, m$, *be given by*

$$\theta_i(dx, dv) := \sum_{k=1}^{N} \alpha_k^i c_k(x) \delta_{v_k(x)} dx. \tag{2.35}$$

We can observe that for $\varphi \in C_c^1(\Omega)$ *and* $i = 1, \dots, d$ *we have*

$$\begin{aligned}\int_{\Omega \times \mathbb{S}^{d-1}} \varphi(x) \nabla c(x_0, x) \cdot v \, \theta_i(dx, dv) &= \lim_{\varepsilon \to 0^+} \sum_{\sigma \in C_i^\varepsilon} \sum_{k=1}^{N(\sigma)-1} |y_k - y_{k-1}|^{d-1} \varphi(y_{k-1})(c^\varepsilon(x_0^\varepsilon, y_k) - c^\varepsilon(x_0^\varepsilon, y_{k-1} \\ &= \lim_{\varepsilon \to 0^+} \sum_{k \,:\, \alpha_k^i = 1} \sum_{(x,e) \in E_k^\varepsilon} |e|^{d-1} \varphi(x)(c^\varepsilon(x_0^\varepsilon, x+e) - c^\varepsilon(x_0^\varepsilon, x)).\end{aligned}$$

In particular, recalling Assumption 2.7, for $\varphi \in C_c^1(\Omega)$ and $k = 1, \dots, N$, we have

$$\begin{aligned}\int_\Omega c_k(x) \varphi(x) \nabla c(x_0, x) \cdot v_k(x) \, dx &= \frac{1}{n_k} \lim_{\varepsilon \to 0^+} \sum_{(x,e) \in E_k^\varepsilon} \sum_{(x_l, e_l) \subset \sigma_{(x,e)}} |e_l|^{d-1} \varphi(x_l)(c^\varepsilon(x_0^\varepsilon, x_l + e_l) - c^\varepsilon(x_0^\varepsilon, x_l) \\ &= \lim_{\varepsilon \to 0^+} \sum_{(x,e) \in E_k^\varepsilon} |e|^{d-1} \varphi(x)(c^\varepsilon(x_0^\varepsilon, x+e) - c^\varepsilon(x_0^\varepsilon, x)).\end{aligned}$$

The last equality comes from Assumption 2.7 and approximations on φ and c^ε.

Lemma 2.8. *Let* $x_0 \in \Omega$ *and* c *be defined by* (2.33), *one has*

1. *For every* $w \in C_c^\infty(\Omega, \mathbb{R}^d)$, *the following inequality holds*

$$\int_\Omega \nabla_x c(x_0, x) \cdot w(x) \, dx \leq \int_\Omega \Phi_\xi(x, w(x)) \, dx. \tag{2.36}$$

2. $c \leq \bar{c}_\xi$ *and so one has the* Γ*-liminf inequality.*

Proof. 1. We still take the notation $\xi_k(x) = \xi(x, v_k(x))$. Let $\alpha(x) = (\alpha_1(x), ..., \alpha_N(x))$ be a minimizing decomposition of $w(x)$ i.e. for all $x \in \Omega$

$$\inf_{X \in A_x^{w(x)}} \sum_{k=1}^{N} x_k \xi_k(x) = \sum_{k=1}^{N} \alpha_k(x) \xi_k(x)$$

with of course $w(x) = \sum_{k=1}^N \alpha_k(x) v_k(x)$ and $\alpha_k(x) \geq 0$. Then we have

$$\int_\Omega \nabla_x c(x_0, x) \cdot w(x) \, dx = \sum_{k=1}^{N} \int_\Omega \alpha_k(x) \, \nabla_x c(x_0, x) \cdot v_k(x) \, dx.$$

However the α_k's are not necessarily smooth so we must regularize the α_k's to pass to the limit. As usual, we consider a mollifying sequence (ρ^δ) (with $\delta > 0$), write

$$\alpha_k^\delta := \rho^\delta \star \alpha_k \;\text{ and }\; w^\delta = \sum_{k=1}^{N} \alpha_k^\delta v_k$$

for $k = 1, \dots, N$. Hence we have

$$\int_\Omega \nabla_x c(x_0, \cdot) \cdot w = \lim_{\delta \to 0^+} \int_\Omega \nabla_x c(x_0, \cdot) \cdot w^\delta.$$

Let $x_0^\varepsilon \in N^\varepsilon$ such that $|x_0 - x_0^\varepsilon| \le \varepsilon$ so that we have the uniform convergence of $c^\varepsilon(x_0^\varepsilon, \cdot)$ to $c(x_0, \cdot)$. Due to Remark 2.3 for every $\varphi \in C_c^1(\Omega)$, we know that for $k = 1, \dots, N$,

$$\int_\Omega c_k(x)\varphi(x)\nabla_x c(x_0, x) \cdot v_k(x)\, dx = \lim_{\varepsilon \to 0^+} \sum_{(x,e)\in E_k^\varepsilon} |e|^d \frac{c^\varepsilon(x_0^\varepsilon, x+e) - c^\varepsilon(x_0^\varepsilon, x)}{|e|} \varphi(x).$$

So we may write for a fixed δ

$$\int_\Omega \nabla_x c(x_0, \cdot) \cdot w^\delta = \lim_{\varepsilon \to 0^+} \sum_{k=1}^N \sum_{(x,e)\in E_k^\varepsilon} |e|^d \frac{c^\varepsilon(x_0^\varepsilon, x+e) - c^\varepsilon(x_0^\varepsilon, x)}{|e|} \frac{\alpha_k^\delta(x)}{c_k(x)}.$$

Since $c^\varepsilon(x_0^\varepsilon, x+e) - c^\varepsilon(x_0^\varepsilon, x) \le |e|\xi^\varepsilon(x, e)$, we obtain

$$\begin{aligned}\int_\Omega \nabla_x c(x_0, \cdot) \cdot w^\delta &\le \lim_{\varepsilon \to 0^+} \sum_{(x,e)\in E_k^\varepsilon} |e|^d \xi^\varepsilon(x, e) \frac{\alpha_k^\delta(x)}{c_k(x)} \\ &= \int_\Omega \alpha_k^\delta \xi_k.\end{aligned}$$

Passing to the limit in $\delta \to 0^+$, we finally get

$$\begin{aligned}\int_\Omega \nabla_x c(x_0, \cdot) \cdot w &\le \sum_{k=1}^N \int_\Omega \alpha_k\, \xi_k \\ &= \int_\Omega \inf_{X \in A_x^{w(x)}} \left(\sum_{k=1}^N x_k \xi(x, v_k(x)) \right) dx.\end{aligned}$$

2. First, using (2.36) with $w = \theta v$ for $v \in C_c^\infty(\Omega, \mathbb{R}^d)$ and an arbitrary scalar function $\theta \in C_c^\infty(\Omega, \mathbb{R}), \theta \ge 0$, we deduce from the homogeneity of $z \mapsto \Phi_\xi(x, z)$ that
$$\nabla_x c(x_0, x) \cdot v(x) \le \Phi_\xi(x, v(x)), \text{ a.e. on } \Omega. \tag{2.37}$$
Now let x be a Lebesgue point of both ξ and $\nabla_x c(x_0, \cdot)$, $u \in \mathbb{S}^{d-1}$ and take $v \in C_c^\infty(\Omega, \mathbb{R}^d)$ such that $v = u$ in some neighbourhood of x. By integrating inequality (2.37) over $B_r(x)$, dividing by its measure and letting $r \to 0^+$ we obtain
$$\nabla_x c(x_0, x) \cdot u \le \Phi_\xi(x, u), \text{ a.e. on } \Omega.$$
From Lemma 2.7 the desired result follows.

□

4.2 The Γ-limsup inequality

Given $\xi \in L_+^p(\theta)$, we now prove the Γ-limsup inequality that is there exists a family $\xi^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}$ such that

$$\xi^\varepsilon \to \xi \text{ and } \limsup_{\varepsilon \to 0^+} J^\varepsilon(\xi^\varepsilon) \le J(\xi). \tag{2.38}$$

We first show (4.25) for ξ continuous and then a density argument will allow us to treat the general case.

Step 1 : The case where ξ is continuous.
For every $\varepsilon > 0, (x,e) \in E^\varepsilon$, write

$$\xi^\varepsilon(x,e) := \xi\left(x, \frac{e}{|e|}\right).$$

We have

$$\|\xi^\varepsilon\|_{\varepsilon,p} \to \|\xi\|_p \text{ and } I_0^\varepsilon(\xi^\varepsilon) \to I_0(\xi) \text{ as } \varepsilon \to 0^+.$$

In particular, for $\varepsilon > 0$ small enough, $\|\xi^\varepsilon\|_{\varepsilon,p} \leq 2\|\xi\|_p$ and $\xi^\varepsilon \to \xi$ in the weak sense of definition 2.2. We can proceed analogously to the construction (2.33) of c for the Γ-liminf. We define c^ε on the whole of $\overline{\Omega} \times \overline{\Omega}$ in a similar way and we also have the uniform convergence of c^ε to some c in $C(\overline{\Omega} \times \overline{\Omega})$ (passing up to a subsequence) and $\liminf_{\varepsilon \to 0^+} I_1^\varepsilon(\xi^\varepsilon) = \int_{\overline{\Omega}\times\overline{\Omega}} c d\gamma$ so that to prove (4.25) it is sufficient to show that $c \geq c_\xi = \overline{c}_\xi$. To justify this inequality it is enough to see that by construction for $(x,y) \in N^\varepsilon \times N^\varepsilon$ one has

$$c^\varepsilon(x,y) = \inf_{\sigma \in C_{x,y}^\varepsilon} \int_0^1 \Psi^\varepsilon(\tilde{\sigma}(t), \dot{\tilde{\sigma}}(t)) \cdot \boldsymbol{\xi}^\varepsilon(\tilde{\sigma}(t)) dt \geq c_\xi(x,y)$$

using the uniform convergence of c^ε to c we indeed obtain $c \geq c_\xi = \overline{c}_\xi$.

Step 2 : the general case where ξ is only $L_+^p(\theta)$.
Let $\xi_n \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ such that

$$\|\xi - \xi_n\|_p + \|c_{\xi_n} - \overline{c}_\xi\|_\infty + |I_0(\xi_n) - I_0(\xi)| \leq \frac{1}{n}$$

and

$$\|\xi_n\|_p \leq 2\|\xi\|_p$$

(existence is given by Lemma 2.3). For every $n > 0$ and $\varepsilon > 0$, thanks to Step 1, there exists $\xi_n^\varepsilon \in \mathbb{R}_+^{\#E^\varepsilon}$ such that $\xi_n^\varepsilon \to \xi_n$. Then there exists a nonincreasing sequence $\varepsilon_n > 0$ converging to 0 such that for every $0 < \varepsilon < \varepsilon_n$ we have

$$|I_0^\varepsilon(\xi_n^\varepsilon) - I_0(\xi_n)| \leq \frac{1}{n},\ I_1^\varepsilon(\xi_n^\varepsilon) \geq I_1(\xi_n) - \frac{1}{n} \text{ and } \|\xi_n^\varepsilon\|_{\varepsilon,p} \leq 2\|\xi_n\|_p.$$

For $\varepsilon > 0$, let $n_\varepsilon := \sup\{n; \varepsilon_n \geq \varepsilon\}$ and $\xi^\varepsilon := \xi_{n_\varepsilon}^\varepsilon$ then we get $\xi^\varepsilon \to \xi$ ($\|\xi^\varepsilon\|_{\varepsilon,p} \leq 2\|\xi_n\|_p \leq 4\|\xi\|_p$) as well as

$$|I_0^\varepsilon(\xi^\varepsilon) - I_0(\xi)| \leq \frac{2}{n_\varepsilon} \to 0 \text{ as } \varepsilon \to 0^+$$

and

$$I_1^\varepsilon(\xi^\varepsilon) \geq I_1(\xi_{n_\varepsilon}) - \frac{1}{n_\varepsilon} = \int_{\Omega\times\Omega} c_{\xi_{n_\varepsilon}} d\gamma - \frac{1}{n_\varepsilon}.$$

Since $c_{\xi_{n_\varepsilon}}$ converges to $\overline{c}_\xi$, we then have

$$\liminf I_1(\xi^\varepsilon) \geq I_1(\varepsilon)$$

which completes the proof.

Chapter 3

Optimality conditions and long-term variant

This chapter is the second part of the paper [67]. It takes the notations, definitions and assumptions in the first part.

1 Optimality conditions and continuous Wardrop equilibria

Now we are interested in finding optimality conditions for the limit problem:

$$\inf_{\xi \in L^p_+(\theta)} J(\xi) := \int_{\Omega \times \mathbb{S}^{d-1}} H(x, v, \xi(x,v))\, \theta(dx, dv) - \int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_\xi \, d\gamma, \tag{3.1}$$

through some dual formulation that can be seen in terms of continuous Wardrop equilibria. More precisely, it is in some sense the continuous version of the discrete minimization problem subject to the mass conservation conditions (2.2)-(2.3). Write

$$\mathcal{L} := \{(\sigma, \rho) : \sigma \in W^{1,\infty}([0,1], \overline{\Omega}), \rho \in \mathcal{P}_\sigma \cap L^1([0,1])^N\},$$

where

$$\mathcal{P}_\sigma := \left\{ \rho : t \in [0,1] \mapsto \rho(t) \in \mathbb{R}^N_+ : \dot\sigma(t) = \sum_{k=1}^N v_k(\sigma(t))\, \rho_k(t) \text{ a.e. } t \right\}.$$

We consider $\mathcal{L}$ as a subset of $C([0,1], \mathbb{R}^d) \times L^1([0,1])^N$ i.e. equipped with the product topology, that on $C([0,1], \mathbb{R}^d)$ being the uniform topology and that on $L^1([0,1])^N$ the weak topology. Slightly abusing notations, let us denote $\mathcal{M}^1_+(\mathcal{L})$ the set of Borel probability measures Q on $C([0,1], \mathbb{R}^d) \times L^1([0,1])^N$ such that $Q(\mathcal{L}) = 1$. For $\sigma \in W^{1,\infty}([0,1], \overline{\Omega})$, let us denote by $\tilde\sigma$ the constant speed reparameterization of σ belonging to $W^{1,\infty}([0,1], \overline{\Omega})$ i.e. for $t \in [0,1]$, $\tilde\sigma(t) := \sigma(s^{-1}(t))$, where

$$s(t) := \frac{1}{l(\sigma)} \int_0^t |\dot\sigma(u)|\, du \text{ with } l(\sigma) := \int_0^1 |\dot\sigma(u)|\, du.$$

Likewise for $\rho \in \mathcal{P}_\sigma \cap L^1([0,1])^N$, let $\tilde\rho$ be the reparameterization of ρ i.e.

$$\tilde\rho_k(t) := \frac{l(\sigma)}{|\dot\sigma(s^{-1}(t))|} \rho_k(s^{-1}(t)), \forall t \in [0,1], k = 1, \ldots, N.$$

We have $\tilde{\rho} \in \mathcal{P}_{\tilde{\sigma}} \cap L^1([0,1])^N$ with $\|\tilde{\rho}\|_{L_1} = \|\rho\|_{L^1}$. Define

$$\widetilde{\mathcal{L}} := \{(\sigma,\rho) \in \mathcal{L} : |\dot{\sigma}| \text{ is constant}\} = \{(\tilde{\sigma},\tilde{\rho}), (\sigma,\rho) \in \mathcal{L}\}.$$

Let $Q \in \mathcal{M}^1_+(\mathcal{L})$, we define $\widetilde{Q} \in \mathcal{M}^1_+(\widetilde{\mathcal{L}})$ as the push forward of Q through the map $(\sigma,\rho) \mapsto (\tilde{\sigma},\tilde{\rho})$. Then let us define the set of probability measures on generalized curves that are consistent with the transport plan γ :

$$\mathcal{Q}(\gamma) := \{Q \in \mathcal{M}^1_+(\mathcal{L}) : (e_0,e_1)_\# Q = \gamma\}, \tag{3.2}$$

where e_0 and e_1 are evaluations at time 0 and 1 and $(e_0,e_1)_\# Q$ is the image measure of Q by (e_0,e_1). Thus $Q \in \mathcal{Q}(\gamma)$ means that

$$\int_{\mathcal{L}} \varphi(\sigma(0),\sigma(1))\, dQ(\sigma,\rho) := \int_{\overline{\Omega}\times\overline{\Omega}} \varphi(x,y)\, d\gamma(x,y),\ \ \forall \varphi \in C(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}).$$

This is the continuous analogue of the mass conservation condition (2.2) since Q plays the same role as the paths-flows in the discrete model. Let us now write the analogue of the arc flows induced by $Q \in \mathcal{Q}(\gamma)$; for $k = 1,\dots,N$ let us define the nonnegative measures on $\overline{\Omega} \times \mathbb{S}^{d-1}$, m^Q_k by

$$\int_{\overline{\Omega}\times\mathbb{S}^{d-1}} \varphi(x,v)\, dm^Q_k(x,v) = \int_{\mathcal{L}} \left(\int_0^1 \varphi(\sigma(t), v_k(\sigma(t)))\rho_k(t) dt \right) dQ(\sigma,\rho),$$

for every $\varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R})$. Then the nonnegative measure on $\overline{\Omega} \times \mathbb{S}^{d-1}$ $m^Q = \sum_{k=1}^N m^Q_k$ may be defined by

$$\int_{\overline{\Omega}\times\mathbb{S}^{d-1}} \xi dm^Q = \int_{\mathcal{L}} L_\xi(\sigma,\rho)\, dQ(\sigma,\rho), \forall \xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+) \tag{3.3}$$

where for every $(\sigma,\rho) \in \mathcal{L}$,

$$L_\xi(\sigma,\rho) = \sum_{k=1}^N \int_0^1 \xi(\sigma(t), v_k(\sigma(t)))\rho_k(t)\, dt = \int_0^1 \boldsymbol{\xi}(\sigma(t)) \cdot \rho(t)\, dt, \tag{3.4}$$

with

$$\boldsymbol{\xi}(\sigma(t)) = (\xi(\sigma(t), v_1(\sigma(t))), \dots, \xi(\sigma(t), v_N(\sigma(t)))).$$

Notice that $L_\xi(\sigma,\rho) = L_\xi(\tilde{\sigma},\tilde{\rho})$ for every $(\sigma,\rho) \in \mathcal{L}$ and so $m^{\tilde{Q}} = m^Q$ for every $Q \in \mathcal{M}^1_+(\mathcal{L})$. The p growth asumption (2.13) on $H(x,v,\cdot)$ can be reformulated by a $q = p/(p-1)$ growth on $G(x,v,\cdot)$. To be more precise, we will assume that $g(x,v,\cdot)$ is continuous, positive and increasing in its last argument (so that $G(x,v,\cdot)$ is strictly convex) such that there exists a and b such that $0 < a \leq b$ and

$$am^{q-1} \leq g(x,v,m) \leq b(m^{q-1}+1)\, \forall\, (x,v,m) \in \overline{\Omega} \times \mathbb{S}^{d-1} \times \mathbb{R}_+, \tag{3.5}$$

with $q \in (1, d/(d-1))$. Then let us define

$$\mathcal{Q}^q(\gamma) := \{Q \in \mathcal{Q}(\gamma)\ :\ m^Q \in L^q(\Omega \times \mathbb{S}^{d-1}, \theta)\} \tag{3.6}$$

and assume

$$\mathcal{Q}^q(\gamma) \neq \emptyset. \tag{3.7}$$

This assumption is satisfied for instance when γ is a discrete probability measure on $\overline{\Omega} \times \overline{\Omega}$ and $q < d/(d-1)$. Indeed, first for $Q \in \mathcal{M}^1_+(W^{1,\infty}([0,1],\overline{\Omega}))$, let us define $i_Q \in \mathcal{M}_+(\overline{\Omega})$ as follows

$$\int_\Omega \varphi \, di_Q := \int_{W^{1,\infty}([0,1],\overline{\Omega})} \left(\int_0^1 \varphi(\sigma(t)) |\dot{\sigma}(t)| dt \right) dQ(\sigma) \text{ for } \varphi \in C(\overline{\Omega}, \mathbb{R}).$$

It follows from [23] that there exists $Q \in \mathcal{M}^1_+(W^{1,\infty}([0,1],\overline{\Omega}))$ such that $(e_0, e_1)_\# Q = \gamma$ and $i_Q \in L^q$. For each curve σ, let $\rho^\sigma \in \mathcal{P}_\sigma$ such that $\sum_k \rho^\sigma_k(t) \leq C|\dot{\sigma}(t)|$ (we have the existence thanks to Assumption 2.4). Then we write $\overline{Q} := (id, \rho^{\cdot})_\# Q$. We obtain $\overline{Q} \in \mathcal{Q}^q(\gamma)$ so that we have proved the existence of such kind of measures.

Let $Q \in \mathcal{Q}^q(\gamma)$ and ξ and $\tilde{\xi}$ be in $C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, we have

$$\begin{aligned} \int_{\mathcal{L}} \left| L_\xi(\sigma,\rho) - L_{\tilde{\xi}}(\sigma,\rho) \right| dQ(\sigma,\rho) &= \int_{\mathcal{L}} \left| \int_0^1 (\boldsymbol{\xi}(\sigma(t)) - \tilde{\boldsymbol{\xi}}(\sigma(t))) \cdot \rho(t) \, dt \right| dQ(\sigma,\rho) \\ &\leq \int_{\Omega \times \mathbb{S}^{d-1}} \left| \xi - \tilde{\xi} \right| m^Q \, \theta(dx,dv) \\ &\leq \|\xi - \tilde{\xi}\|_{L^p(\theta)} \|m^Q\|_{L^q(\theta)}. \end{aligned}$$

So if $\xi \in L^p_+(\theta)$ and $(\xi_n)_n$ is a sequence in $C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$ that converges in $L^p(\theta)$ to ξ then L_{ξ_n} is a Cauchy sequence in $L^1(\mathcal{L}, Q)$ and its limit (that we continue to denote by L_ξ) does not depend on the approximating sequence $(\xi_n)_n$. This suggests us to define L_ξ in an $L^1(\mathcal{L}, Q)$ sense for every $\xi \in L^p_+(\theta)$ and $Q \in \mathcal{Q}^q(\gamma)$. For every $\xi \in L^p_+(\theta)$ and $Q \in \mathcal{Q}^q(\gamma)$, by proceeding as for Lemma 3.6 in [39], we have

$$\int_{\Omega \times \mathbb{S}^{d-1}} \xi \cdot m^Q \, \theta(dx,dv) = \int_{\mathcal{L}} L_\xi(\sigma,\rho) \, dQ(\sigma,\rho), \tag{3.8}$$

and

$$\overline{c}_\xi(\sigma(0),\sigma(1)) \leq L_\xi(\sigma,\rho) \text{ for } Q-\text{a.e. } (\sigma,\rho) \in \mathcal{L}. \tag{3.9}$$

Hence using the fact that $Q \in \mathcal{Q}^q(\gamma)$ and (3.8)-(3.9), we obtain

$$\int_{\overline{\Omega} \times \overline{\Omega}} \overline{c}_\xi d\gamma = \int_{\mathcal{L}} \overline{c}_\xi(\sigma(0),\sigma(1)) \, dQ(\sigma,\rho) \leq \int_{\Omega \times \mathbb{S}^{d-1}} \xi \cdot m^Q. \tag{3.10}$$

Let $\xi \in L^p_+(\theta)$ and $Q \in \mathcal{Q}^q(\gamma)$, it follows from Young's inequality that

$$\begin{aligned} &\int_{\Omega \times \mathbb{S}^{d-1}} H(x,v,\xi(x,v)) \, \theta(dx,dv) \\ &\qquad \geq \int_{\Omega \times \mathbb{S}^{d-1}} \xi \cdot m^Q \, \theta(dx,dv) - \int_{\Omega \times \mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right) \, \theta(dx,dv) \end{aligned} \tag{3.11}$$

so that we have

$$\inf_{\xi \in L^p_+(\theta)} J(\xi) \geq \sup_{Q \in \mathcal{Q}^q(\gamma)} - \int_{\Omega \times \mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right) \, \theta(dx,dv). \tag{3.12}$$

The dual formulation of (3.1) then is

$$\sup_{Q \in \mathcal{Q}^q(\gamma)} - \int_{\Omega \times \mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right) \, \theta(dx,dv). \tag{3.13}$$

We can note the analogy between (3.13) and the discrete problem that consists in minimizing (2.4) subject to the mass conservation conditions (2.2)-(2.3). Then we establish the following theorem, that specifies relations between (3.13) and (3.1) and that gives the connection with Wardrop equilibria:

Theorem 3.1. *Under assumptions* (3.5) *and* (3.7), *we have:*

1. *The problem* (3.13) *admits solutions.*
2. $\overline{Q} \in \mathcal{Q}^q(\gamma)$ *solves* (3.13) *if and only if*

$$\int_{\mathcal{L}} L_{\xi_{\overline{Q}}}(\sigma,\rho)\, d\overline{Q}(\sigma,\rho) = \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0),\sigma(1))\, d\overline{Q}(\sigma,\rho) \tag{3.14}$$

where $\xi_{\overline{Q}}(x,v) := g\left(x,v,m^{\overline{Q}}(x,v)\right)$.

3. *Equality holds :* $\inf(3.1) = \sup(3.13)$. *Moreover if* $\overline{Q}$ *solves* (3.13) *then* $\xi_{\overline{Q}}$ *solves* (3.1).

It is the main result of this section. To prove it, we need some lemmas. First, let us start with a preliminary lemma on the v_k's that is a consequence of Assumption 2.4.

Lemma 3.1. *For all subset* $I \subset \{1,\dots,N\}$, *we are in one of the two following cases :*

1. $0 \in Conv(\{v_i(x)\}_{i\in I})$ *for every* $x \in \overline{\Omega}$,
2. $0 \notin Conv(\{v_i(x)\}_{i\in I})$ *for every* $x \in \overline{\Omega}$.

Moreover, there exists a constant $0 < \delta < 1$ *such that for all subset* $I \subset \{1,\dots,N\}$ *that is in the second case, there exists* $u_x \in Conv(\{v_i(x)\}_{i\in I}$ *for all* $x \in \overline{\Omega}$ *such that*

$$v_i(x) \cdot \frac{u_x}{|u_x|} \geq \delta \textit{ for all } i \in I.$$

Proof. We will use the fact that $\overline{\Omega}$ is connected. The first property is obviously closed since the v_k's are continuous. Let us now show that the second one is closed. Let $I \subset \{1,\dots,N\}$, assume by contradiction that there exists a sequence $\{x_n\}_{n\geq 0} \subset \overline{\Omega}$ converging to $x \in \overline{\Omega}$ such that $0 \notin C_n = \mathrm{Conv}(\{v_i(x_n)\}_{i\in I})$ for every $n \geq 0$ and $0 \in C = \mathrm{Conv}(\{v_i(x)\}_{i\in I})$. So there exists $\{\lambda_i\}_{i\in I}$ such that $\sum_{i\in I} \lambda_i v_i(x) = 0, \lambda_i \geq 0$ and $\sum_{i\in I}\lambda_i = 1$. Without loss of generality, we can assume that the λ_i's are positive. Then we have that $v^n = \sum_{i\in I}\lambda_i v_i(x_n) \neq 0$ and converges to 0 as $n \to +\infty$. Let $\beta_n > 0$ such that $|\beta_n v^n| = 1$ then β_n converges to $+\infty$. Thanks to Assumption 2.4, with $\xi = (\xi_1,\dots,\xi_N), \xi_i = 0$ if $i \in I, 1$ otherwise, there exists $\{z_i^n\}_{i\in I} \subset \mathbb{R}_+^{\#I}$ such that $|z_i^n| \leq C$ and $\sum_{i\in I} z_i^n v_i(x_n) = \beta_n v^n$ for all $i \in I$ and $n \geq 0$. Then we obtain that $\sum_{i\in I}(\beta_n\lambda_i - z_i^n)v_i(x_n) = 0$. But for n large enough, we have that $\beta_n\lambda_i - z_i^n > 0$ for every $i \in I$, which is a contradiction.

For a subset $I \subset \{1,\dots,N\}$ that is in the second case, $\mathrm{Conv}(\{v_i(x)\}_{i\in I})$ is contained in a salient (pointed) cone for all $x \in \overline{\Omega}$. Let us recall that a set A is a salient cone if and only if $A \cap (-A) \subseteq \{0\}$, that is, if x is in $A\backslash\{0\}$ then $-x$ is not in A. For all $x \in \overline{\Omega}$, we can think of u_x as being in the medial axis of this cone. Since the v_k's are continuous and $\overline{\Omega}$ is compact, we have the desired result. □

Now let us define the following sets

$$\mathcal{L}^C := \left\{(\sigma,\rho) \in \mathcal{L} : \sum_{k=1}^N \rho_k(t) \leq C|\dot{\sigma}(t)| \text{ a.e. } t \in [0,1]\right\}$$

for some constants $C > 0$. Now let us notice that we can simplify the problem (3.13) with the following lemma:

Lemma 3.2. *For a well-chosen constant $C' > 1$, one has*

$$\inf_{Q\in\mathcal{Q}^q(\gamma)} \int_{\Omega\times\mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right)\,\theta(dx,dv)$$
$$= \inf_{Q\in\mathcal{Q}^q(\gamma)} \left\{\int_{\Omega\times\mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right)\,\theta(dx,dv) : Q\left(\mathcal{L}^{C'}\right)=1\right\}.$$

Proof. We set $C' = 1/\delta$ where δ is given by Lemma 3.1. Let $(\sigma,\rho)\in\mathcal{L}$. We will prove that there exists $\overline{\rho}\in\mathcal{P}_\sigma$ such that for all $t\in[0,1]$, $\overline{\rho}_k(t)\le\rho_k(t)$ for all $k=1,\dots,N$ and $\sum_{k=1}^N\overline{\rho}_k(t)\le C'|\dot{\sigma}(t)|$. Let $t\in[0,1]$ such that $\sum_{k=1}^N\rho_k(t) > C'|\dot\sigma(t)|$. Let us denote I the subset of $\{1,\dots,N\}$ such that for every $k\in I, \rho_k(t)>0$. First, if $0\in \mathrm{Conv}(\{v_k(\sigma(t))\}_{k\in I})$, there exists a conical combination of 0

$$\sum_{k\in I}\lambda_k v_k(\sigma(t)) = 0$$

with the λ_k's ≥ 0. Then we write $\overline{\rho}_k(t)=\rho_k(t)-\lambda\lambda_k$ (we take $\lambda_k=0$ for $k\notin I$) where

$$\lambda := \min_{k\in I:\lambda_k\neq 0}\left\{\frac{\rho_k(t)}{\lambda_k}\right\}.$$

We set $\bar I$ the subset of I such that for every $k\in\bar I, \overline{\rho}_k(t)>0$. We restart with $\overline{\rho}$ and we continue until $0\notin\mathrm{Conv}(\{v_k(\sigma(t))\}_{k\in\bar I})$. Let u be as in Lemma 3.1 for $I=\bar I$ and $x=\sigma(t)$ with $|u|=1$ (take $u/|u|$ if necessarily). Then we have:

$$|\dot\sigma(t)|\ge\dot\sigma(t)\cdot u=\sum_{k=1}^N\overline{\rho}_k(t)v_k(\sigma(t))\cdot u\ge\delta\sum_{k=1}^N\overline{\rho}_k(t)$$

so that $\sum_{k=1}^N\overline{\rho}_k(t)\le C'|\dot\sigma(t)|$. For $Q\in\mathcal{M}_+^1(\mathcal{L})$, we denote by $\overline{Q}\in\mathcal{M}_+^1(\mathcal{L}^{C'})$ the push forward of Q through the map $(\sigma,\rho)\mapsto(\sigma,\overline{\rho})$. Then we have $m^{\overline{Q}}\le m^Q$. Since $G(m,v,\cdot)$ is nondecreasing, we have:

$$\int_{\Omega\times\mathbb{S}^{d-1}} G\left(x,v,m^{\overline{Q}}(x,v)\right)\,\theta(dx,dv)\le\int_{\Omega\times\mathbb{S}^{d-1}} G\left(x,v,m^Q(x,v)\right)\,\theta(dx,dv).$$

□

To prove that the problem (3.13) has solutions, a natural idea would be to take a maximizing sequence $\{Q_n\}_{n\ge0}$ for (3.13) and to show that it converges to $Q\in\mathcal{Q}^q(\gamma)$ that solves (3.13). For this, we would like to use Prokhorov's theorem which would allow us to obtain the tightness of $\{\widetilde{Q}_n\}$ and $\star$-weak convergence of $\{Q_n\}$ to a measure in $\mathcal{M}_+^1(\mathcal{L})$. Unfortunately, the space $C([0,1],\mathbb{R}^d)\times L^1([0,1])^N$ is not a Polish space for the considered topology (because of the weak topology of $L^1([0,1])$). So we will work with Young's measures in order to apply Prokhorov's theorem. Let us define the set

$$\mathcal{U}:=C([0,1],\mathbb{R}^d)\times\mathfrak{P}_1(\mathbb{R}^d\times[0,1]) \tag{3.15}$$

where for a Polish space (E,d), we set

$$\mathfrak{P}_1(E):=\left\{\mu\in\mathcal{M}_+^1(E):\int_E \mathrm{d}(x,x')\,d\mu(x)<+\infty\ \text{ for some } x'\in E\right\}.$$

We equip $\mathcal{U}$ with the product topology, that on $C([0,1],\mathbb{R}^d)$ being the uniform topology and $\mathfrak{P}_1(\mathbb{R}^d \times [0,1])$ being endowed with the 1-Wasserstein distance

$$W_1(\mu,\nu) := \min\left\{\int_{E^2} \mathrm{d}(x_1,x_2)\, d\pi(x_1,x_2) : \pi \in \Pi(\mu,\nu)\right\}$$

where $E = \mathbb{R}^d \times [0,1]$, d is the usual distance on E, $(\mu,\nu) \in \mathfrak{P}_1(E)^2$ and $\Pi(\mu,\nu)$ is the set of transport plans between μ and ν, that is, the set of probability measures π on E^2, having μ and ν as marginals:

$$\int_{E\times E} \varphi(x) d\pi(x,y) = \int_E \varphi(x) d\mu(x) \text{ and } \int_{E\times E} \varphi(y) d\pi(x,y) = \int_E \varphi(x) d\nu(x), \quad (3.16)$$

for every $\varphi \in C(E,\mathbb{R})$. The set $\mathcal{U}$ is a Polish space (see [9]). Let us denote by λ the Lebesgue measure on $[0,1]$ and let us consider the subset $\mathcal{S}$ of $\mathcal{U}$:

$$\mathcal{S} := \left\{(\sigma, \nu_t \otimes \lambda) : \sigma \in W^{1,\infty}([0,1],\overline{\Omega}), \nu_t \otimes \lambda \in \mathfrak{P}_1(E), \nu_t \in \mathfrak{M}^t_\sigma \text{ a.e. } t\right\},$$

where for $t \in [0,1]$ and $\sigma \in W^{1,\infty}([0,1],\overline{\Omega})$,

$$\mathfrak{M}^t_\sigma := \left\{\nu_t \in \mathcal{M}^1_+(\mathbb{R}^d) : \text{supp } \nu_t \subset \bigcup_{k=1}^N \mathbb{R}_+ v_k(\sigma(t)) \text{ and } \dot\sigma(t) = \int_{\mathbb{R}^d} v\, d\nu_t(v)\right\}.$$

Here the set $\mathbb{R}_+ v_k(\sigma(t))$ is the half-line $\{a v_k(\sigma(t)) : a \in \mathbb{R}_+\}$. The Young measures $\nu_t \otimes \lambda$ are the analogue of the decompositions $\rho \in \mathcal{P}_\sigma$. For the general theory of the Young measures, see for instance [85].

Let us define the set of probability measures on curves $(\sigma, \nu_t \otimes \lambda)$ that are consistent with the transport plan γ :

$$\mathcal{X}(\gamma) := \{X \in \mathcal{M}^1_+(\mathcal{S}) : (e_0, e_1)_\# X = \gamma\}. \quad (3.17)$$

This is the analogue of (3.2). Let us now write the analogue of m^Q (given by (4.9)) as follows:

$$\int_{\overline{\Omega}\times\mathbb{S}^{d-1}} \xi di^X := \int_{\mathcal{S}} \overline{L}_\xi(\sigma,\kappa)\, dX(\sigma,\kappa), \forall \xi \in C(\overline{\Omega}\times\mathbb{S}^{d-1},\mathbb{R}) \quad (3.18)$$

where for every $(\sigma, \kappa = \nu_t \otimes \lambda) \in \mathcal{L}$,

$$\overline{L}_\xi(\sigma,\kappa) := \int_0^1 \left(\int_{\mathbb{R}^d} \xi\left(\sigma(t), \frac{v}{|v|}\right) |v|\, d\nu_t(v)\right) dt. \quad (3.19)$$

Then let us define

$$\mathcal{X}^q(\gamma) := \{X \in \mathcal{X}(\gamma) \ : \ i^X \in L^q(\Omega\times\mathbb{S}^{d-1},\theta)\} \quad (3.20)$$

Let $X \in \mathcal{X}^q(\gamma)$. By the same reasoning as for $Q \in \mathcal{Q}^q(\gamma)$, if $\xi \in L^p_+(\theta)$, we denote by $\overline{L}_\xi$ the limit of the Cauchy sequence $\overline{L}_{\xi_n}$ in $L^1(\mathcal{S},X)$ for any sequence $(\xi_n)_n$ converging in $L^p(\theta)$ to ξ. We may write the analogue of the problem (3.13) :

$$\sup_{X\in\mathcal{X}^q(\gamma)} -\int_{\Omega\times\mathbb{S}^{d-1}} G\left(x,v,i^X(x,v)\right)\, \theta(dx,dv). \quad (3.21)$$

Lemma 3.3. *One has* sup (3.13) = sup (3.21).

Proof. Let $Q \in \mathcal{Q}^q(\gamma)$ and $\sigma \in W^{1,\infty}([0,1],\overline{\Omega})$. For $\rho = (\rho_1, \dots, \rho_N) \in \mathcal{P}_\sigma$, we define the Young's measure $\nu_t^\rho \otimes \lambda$ as follows :

$$\nu_t^\rho := \sum_{k=1}^N \frac{\rho_k(t)}{|\rho(t)|_1} \delta_{\{|\rho(t)|_1 v_k(\sigma(t))\}},$$

where $|\rho(t)|_1 = \sum_{k=1}^N \rho_k(t)$ for every $t \in [0,1]$. We consider the measure $X^Q \in \mathcal{X}^q(\gamma)$ defined by

$$\int_{\mathcal{S}} \varphi \, dX^Q := \int_{\mathcal{L}} \varphi(\sigma, \nu_t^\rho \otimes \lambda) \, dQ(\sigma, \rho), \text{ for all } \varphi \in C(\mathcal{S}, \mathbb{R}).$$

Since we have $m^Q = i^{X^Q}$ we immediately get sup (3.13) $\leq$ sup (3.21).

For the converse inequality, let $X \in \mathcal{X}^q(\gamma)$, we build $Q^X \in \mathcal{Q}^q(\gamma)$. Let $(\sigma, \nu_t \otimes \lambda) \in \mathcal{S}$, recalling that one has supp $\nu_t \subset \bigcup_{k=1}^N \mathbb{R}_+ v_k(\sigma(t))$ for $t \in [0,1]$, we define $\rho^\nu \in \mathcal{P}_\sigma$ as follows

$$\rho_k^\nu(t) := \int_{\mathbb{R}_+ v_k(\sigma(t))} |v| \, d\nu_t(v), \text{ for all } k = 1, \dots, N$$

and $\rho^\nu = (\rho_1^\nu, \dots, \rho_N^\nu)$ if the $v_k(\sigma(t))$'s are pairwise distinct. Otherwise, let us decompose $\{1, \dots, N\} = \bigcup_{j=1}^s I_j$ where the I_k's are pairwise disjoint and such that for all $j = 1, \dots, s$ and $k \in I_j, v_k(\sigma(t)) = v_j$ where the v_j's are pairwise distinct. Then for all $j = 1, \dots, s$ and $k \in I_j$, we set

$$\rho_k^\nu(t) := \frac{1}{\# I_j} \int_{\mathbb{R}_+ v_j} |v| \, d\nu_t(v).$$

The element ρ^ν is in $\mathcal{P}_\sigma$. Similarly, we set

$$\int_{\mathcal{L}} \varphi \, dQ^X := \int_{\mathcal{S}} \varphi(\sigma, \rho^\nu) \, dQ(\sigma, \nu_t \otimes \lambda) \text{ for all } \varphi \in C(\mathcal{L}, \mathbb{R}).$$

From the fact that $m^{Q^X} = i^X$ it follows that sup (3.13) $\geq$ sup (3.21). □

Let us notice that with the previous proof for $(\sigma, \nu_t \otimes \lambda) \in \mathcal{S}$, we may build $\widetilde{\nu}$ as a sum of Dirac measures :

$$\widetilde{\nu}_t := \sum_{k=1}^N \frac{\rho_k^\nu(t)}{|\rho^\nu(t)|_1} \delta_{\{|\rho^\nu(t)|_1 v_k(\sigma(t))\}}$$

where ρ^ν is given in the previous proof. Therefore it follows from the same reasoning as in the proof of Lemma 3.2 that we may take $\rho^\nu \in \mathcal{P}_\sigma$ such that for $t \in [0,1], \sum_{k=1}^N \rho_k(t) \leq C'|\dot{\sigma}(t)|$. Moreover, we can choose (σ, ρ) only in $\widetilde{\mathcal{L}}$ with $|\dot{\sigma}|$ constant. Then the new measure $\sum_{k=1}^N \frac{\rho_k^\nu(t)}{|\rho^\nu(t)|_1} \delta_{\{|\rho(t)|_1 v_k(\sigma(t))\}}$ that we continue to denote by $\widetilde{\nu}_t$ by abuse of notations is in $\mathfrak{M}_\sigma^t$. Let us define

$$\mathcal{S}^{C'} := \{(\sigma, \nu_t \otimes \lambda) \in \mathcal{S} : \text{supp } \nu_t \cap \mathbb{R}_+ v_k(\sigma(t)) = \{\rho(t) v_k(\sigma(t))\}$$
$$\text{with } \rho(t) \leq C'|\dot{\sigma}(t)| \text{ for } k = 1, \dots, N \text{ and } t \in [0,1]\}$$

and

$$\widetilde{\mathcal{S}} := \{(\sigma, \nu_t \otimes \lambda) \in \mathcal{S} : |\dot{\sigma}| \text{ is constant}\}.$$

For $X \in \mathcal{M}_+^1(\mathcal{S})$, we denote by $\widetilde{X} \in \mathcal{M}_+^1(\mathcal{S}^{C'} \cap \widetilde{\mathcal{S}})$ the push forward of X through the map $(\sigma, \nu_t \otimes \lambda) \mapsto (\tilde{\sigma}, \tilde{\nu}_t \otimes \lambda)$. Then we have $i^{\widetilde{X}} \leq i^X$. Since $G(m, v, \cdot)$ is nondecreasing, we may consider only the measures $\widetilde{X} \in \mathcal{M}_+^1(\mathcal{S}^{C'} \cap \widetilde{\mathcal{S}})$ for the problem (3.21).

We now adapt the proof in [12]. In particular we have to generalize Lemmas 2.7 and 2.8 in [39], this becomes

Lemma 3.4. *For every $\varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, $\overline{L}_\varphi$ is l.s.c. on $\mathcal{S}$ for the topology defined above.*

Proof. Let $(\sigma, \nu_t \otimes \lambda) \in \mathcal{S}$ and $(\sigma^n, \nu_t^n \otimes \lambda)$ be a sequence converging to $(\sigma, \nu_t \otimes \lambda) \in \mathcal{S}$. Then by definition, we have

$$\overline{L}_\varphi(\sigma^n, \nu_t^n \otimes \lambda) = \int_0^1 \left(\int_{\mathbb{R}^d} \varphi\left(\sigma^n(t), \frac{v}{|v|}\right) |v| \, d\nu_t^n(v) \right) dt.$$

We have $\sigma^n \to \sigma$ in $C([0,1])$ so that $\varphi(\sigma^n(\cdot), \frac{v}{|v|})$ converges strongly in L^∞. Since ν_t^n narrowly converges to ν_t for almost $t \in [0,1]$ and the function $(t, v) \mapsto \varphi(\sigma(t), \frac{v}{|v|})|v|$ is the upper limit of $(t, v) \mapsto \varphi(\sigma(t), \frac{v}{|v|}) \min(|v|, n)$, that is continuous and bounded, as $n \to +\infty$, we obtain the desired result. $\square$

Lemma 3.5. *Let $(X_n)_n \in \mathcal{M}_+^1(\mathcal{U})^{\mathbb{N}}$ be such that $X_n(\mathcal{S}^{C'}) = 1$ for every n and there exists a constant $M > 0$ such that*

$$\sup_n \int_{\mathcal{S}} l(\sigma) \, dX_n(\sigma, \nu_t \otimes \lambda) \leq M.$$

Then the sequence $(\widetilde{X}_n)_n$ is tight and admits a subsequence that weakly-$\star$ converges to a probability measure X such that $X(\mathcal{S}) = 1$.

Proof. For every $K > 0$, let us define the following subset of $\widetilde{\mathcal{S}}^{C'}$

$$\widetilde{\mathcal{S}}_K := \left\{ (\sigma, \nu_t \otimes \lambda) \in \widetilde{\mathcal{S}} : |\dot{\sigma}| \leq K \text{ and supp } \nu_t \leq B_{C'K} \right\}$$

where C' is the constant given by Lemma 3.2. Let us show that $\widetilde{\mathcal{S}}_K$ is relatively compact in $\mathcal{S}$. First, the set $\{\sigma \in W^{1,\infty}([0,1], \overline{\Omega}) : \sigma \; K\text{-Lipschitz continuous}\}$ is compact in $C([0,1], \overline{\Omega})$ thanks to Ascoli's theorem. The set of probability measures with support in $B_{C'K}$ is compact due to the Banach-Alaoglu-Bourbaki theorem. Let a sequence $(\sigma^n, \nu_t^n \otimes \lambda) \in (\widetilde{\mathcal{S}}_K)^{\mathbb{N}}$ converging to $(\sigma, \nu_t \otimes \lambda) \in \mathcal{S}$, prove that $(\sigma, \nu_t \otimes \lambda) \in \widetilde{\mathcal{S}}_K$.

1.) supp $\nu_t \subset \bigcup_{k=1}^N \mathbb{R}_+ v_k(\sigma(t))$.

First let us show that the function $\varphi : (x, v) \mapsto \text{dist}(v, \bigcup_{k=1}^N \mathbb{R}_+ v_k(x))$ is continuous on $\mathbb{R}^d \times \overline{\Omega}$. Let $(x, v) \in \overline{\Omega} \times \mathbb{R}^d$ and some sequences x^n in $\overline{\Omega}$, respectively v^n in $\mathbb{R}^d$ converging to x, respectively v. Then there exists some constants $0 \leq \lambda_k \leq |v|$ and $0 \leq \lambda_k^n \leq |v^n|$ such that $\lambda_k v_k(x)$, respectively $\lambda_k^n v_k(x^n)$, is the projection of v, respectively v^n, on $\mathbb{R}_+ v_k(x)$, respectively $\mathbb{R}_+ v_k(x^n)$ for $k = 1, \ldots, N$ and $n \in \mathbb{N}$. Then we have

$$\begin{aligned}|\varphi(x,v) - \varphi(x^n, v^n)| &\leq \sum_{k=1}^{N} |\mathrm{dist}(v, \mathbb{R}_+ v_k(x)) - \mathrm{dist}(v^n, \mathbb{R}_+ v_k(x^n))| \\ &= \sum_{k=1}^{N} ||v - \lambda_k v_k(x)| - |v^n - \lambda_k^n v_k(x^n)|| \\ &\leq \sum_{k=1}^{N} (|v - v^n| + |\lambda_k v_k(x) - \lambda_k^n v_k(x^n)|) \\ &\longrightarrow 0 \text{ as } n \to +\infty\end{aligned}$$

since the v_k's are continuous and so the sequence λ_k^n converges to λ_k for $k = 1, \ldots, N$. We then have :

$$\begin{aligned}\int_0^1 \left(\int_{\mathbb{R}^d} \varphi(v, \sigma(t))\, d\nu_t(v) \right) dt &= \int_0^1 \left(\int_{B_{C'K}} \varphi(v, \sigma(t))\, d\nu_t(v) \right) dt \\ &= \lim_{n \to +\infty} \int_0^1 \left(\int_{B_{C'K}} \varphi(v, \sigma^n(t))\, d\nu_t^n(v) \right) dt \\ &= 0.\end{aligned}$$

So $\varphi(x,v) = 0$ $d\nu_t \otimes dt$-a.e. and the support of ν_t is in $\bigcup_{k=1}^N \mathbb{R}_+ v_k(\sigma(t))$ for $t \in [0,1]$.

2.) $\dot\sigma(t) = \int_{\mathbb{R}^d} v\, d\nu_t(v)$.

By definition, for $n \geq 0$ and $(s,t) \in [0,1]^2$, we have :

$$\sigma^n(t) - \sigma^n(s) = \int_s^t \int_{B_{C'K}} v\, d\nu_t^n(v) \otimes \lambda = (v \mathbf{1}_{B_{C'K}} \otimes \mathbf{1}_{[s,t]}; \nu_t^n \otimes \lambda).$$

Obviously, the sequence $\{\sigma^n(t) - \sigma^n(s)\}_{n \geq 0}$ converges to $\sigma(t) - \sigma(s)$ (since σ^n uniformly converges to σ). For the term in the right-hand side, it is sufficient to take a sequence $\{\Phi_\varepsilon\}_{\varepsilon > 0}$ in $C_b(\mathbb{R}^d \times [0,1])^{\mathbb{N}}$ converging to $(v,t) \mapsto v \mathbf{1}_{B_{C'K}} \otimes \mathbf{1}_{[s,t]}$ in $L^1(\mathbb{R}^d \times [0,1])$ as $\varepsilon \to 0^+$.

Now let us justify the tightness of $(\widetilde{X}_n)_n$:

$$\begin{aligned}\widetilde{X}_n\left((\widetilde{\mathcal{S}}_K)^c \right) &\leq \widetilde{X}_n\left(\left\{ (\sigma, \nu_t \otimes \lambda) \in \widetilde{\mathcal{S}} \cap \mathcal{S}^{C'} : |\dot\sigma| > K \right\} \right) \\ &\quad + \widetilde{X}_n\left(\left\{ (\sigma, \nu_t \otimes \lambda) \in \widetilde{\mathcal{S}} \cap \mathcal{S}^{C'} : \mathrm{supp}(\nu_t) \not\subseteq B_{C'K} \right\} \right) \\ &\leq 2\widetilde{X}_n\left(\left\{ (\sigma, \nu_t \otimes \lambda) \in \widetilde{\mathcal{S}} \cap \mathcal{S}^{C'} : |\dot\sigma| > K \right\} \right) \\ &\leq 2X_n\left(\{ (\sigma, \nu_t \otimes \lambda) \in \mathcal{S} : l(\sigma) > K \} \right) \\ &\leq \frac{2}{K} \int_{\mathcal{S}} l(\sigma)\, dX_n(\sigma, \nu_t \otimes \lambda) \\ &\leq 2\frac{M}{K} \to 0 \text{ as } K \to +\infty.\end{aligned}$$

Due to Prokhorov's theorem we can then assume that passing up to a subsequence, $(\widetilde{X}_n)_n$ weakly-$\star$ converges to $X \in \mathcal{M}_+^1(\mathcal{U})$. It remains to show that $X(\mathcal{S}) = 1$. For $K > 0$, let us define the closed set

$$\mathcal{S}_K := \{ (\sigma, \nu_t \otimes \lambda) \in \mathcal{S} : l(\sigma) \leq K \text{ and supp } \nu_t \subset B_{C'K} \}.$$

It follows from the previous computation, the fact that the measures $\widetilde{X}_n$ are concentrated on $\widetilde{\mathcal{S}}$ and Portmanteau's theorem that

$$\begin{aligned}1 = \limsup_n \widetilde{X}_n(\mathcal{S}) &\leq \limsup_n \widetilde{X}_n(\mathcal{S}_K) + \limsup_n \widetilde{X}_n(\mathcal{S}\backslash\mathcal{S}_K)\\ &\leq X(\mathcal{S}_K) + \frac{M}{K}.\end{aligned}$$

Letting K tend to ∞, we then obtain $X(\mathcal{S}) = \sup_K X(\mathcal{S}_K) = 1$. □

Lemma 3.6. *Let $(X_n)_n$ be a sequence in $\mathcal{M}^1_+(\mathcal{S})$ that weakly star converges to some $X \in \mathcal{M}^1_+(\mathcal{S})$. If there exists $i \in \mathcal{M}_+(\overline{\Omega} \times \mathbb{S}^{d-1})$ such that i^{X_n} weakly-$\star$ converges to i in $\mathcal{M}_+(\overline{\Omega} \times \mathbb{S}^{d-1})$ then we have $i^X \leq i$.*

Proof. Let $\varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$. First we obtain

$$\int_{\overline{\Omega}} \varphi \, di = \lim_n \int_{\overline{\Omega}} \varphi di^{X_n} = \lim_n \int_{\mathcal{S}} \overline{L}_\varphi dX_n.$$

Moreover thanks to Lemma 3.4, $X \mapsto \int_{\mathcal{S}} \overline{L}_\varphi dX$ is l.s.c. for the weak star topology of $\mathcal{M}^1_+(\mathcal{S})$. We then obtain

$$\int_{\overline{\Omega}} \varphi di \geq \int_{\mathcal{S}} \overline{L}_\varphi dX = \int_{\overline{\Omega}} \varphi di^X.$$

□

Proof. (of Theorem 3.1)

1. Thanks to Lemma 3.3, we consider the problem (3.21). Due to (3.5) the value of problem (3.21) is finite. Let $(X_n)_n$ be a maximizing sequence of (3.21). Since $i^X \geq i^{\widetilde{X}}$, we can assume $X_n = \widetilde{X}_n$ for all n. Still from (3.5) it follows that $(i^{X_n})_n$ is bounded in $L^q(\theta)$. So, passing up to a subsequence, we can assume that $(i^{X_n})_n$ weakly converges in $L^q(\theta)$ to some i. Moreover, since $(i^{X_n})_n$ is bounded in $L^q(\theta)$ so in $L^1(\theta)$, we have

$$\begin{aligned}\sup_n \int_{\mathcal{S}} l(\sigma) \, dX_n(\sigma, \nu_t \otimes \lambda) &\leq \sup_n \int_{\mathcal{S}} \left(\int_0^1 \left(\int_{\mathbb{R}^d} |v| d\nu_t(v)\right) dt\right) dX_n(\sigma, \nu_t \otimes \lambda)\\ &= \sup_n \int_{\Omega \times \mathbb{S}^{d-1}} di^{X_n} < +\infty.\end{aligned}$$

Since $X_n = \widetilde{X}_n$, we can deduce from Lemma 3.5 that, up to a subsequence, $(X_n)_n$ weakly-$\star$ converges to some $X \in \mathcal{M}^1_+(\mathcal{S})$. Using the fact that $\mathcal{X}(\gamma)$ is weakly closed, we see that $X \in \mathcal{X}(\gamma)$ and Lemma 3.6 then imply that $i^X \leq i$ so that $X \in \mathcal{X}^q(\gamma)$. Since $G(x, v, \cdot)$ is convex and nondecreasing, we then have

$$\begin{aligned}\int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, i^X(x, v)\, \theta(dx, dv) &\leq \int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, i(x, v))\, \theta(dx, dv)\\ &\leq \liminf_n \int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, i^{X_n}(x, v))\, \theta(dx, dv),\end{aligned}$$

which proves that X solves (3.21). Thus as mentioned in the proof of Lemma 3.3, there exists $Q \in \mathcal{Q}^q(\gamma)$ such that $m^Q = i^X$ and so Q is a solution of (3.13).

2. First assume that $\overline{Q} \in \mathcal{Q}^q(\gamma)$ satisfies (3.14). Let $Q \in \mathcal{Q}^q(\gamma)$ then by convexity of $G(x, v, \cdot)$, (3.8) and (3.14), we have

$$\begin{aligned}
\int_{\Omega\times\mathbb{S}^{d-1}} & G(x,v,m^Q(x,v))\,\theta(dx,dv) - \int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m^{\overline{Q}}(x,v))\,\theta(dx,dv) \\
&\geq \int_{\Omega\times\mathbb{S}^{d-1}} \xi_{\overline{Q}} \cdot (m^Q - m^{\overline{Q}}) \\
&= \int_{\mathcal{L}} L_{\xi_{\overline{Q}}}(\sigma,\rho)\,dQ(\sigma,\rho) - \int_{\mathcal{L}} L_{\xi_{\overline{Q}}}(\sigma,\rho)\,d\overline{Q}(\sigma,\rho) \\
&\geq \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0),\sigma(1))\,dQ(\sigma,\rho) - \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0),\sigma(1))\,d\overline{Q}(\sigma,\rho) \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\xi_{\overline{Q}}} d\gamma - \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\xi_{\overline{Q}}} d\gamma = 0.
\end{aligned}$$

Therefore $\overline{Q}$ solves (3.13). Now assume that $\overline{Q} \in \mathcal{Q}^q(\gamma)$ solves (3.13), let $Q \in \mathcal{Q}^q(\gamma)$ and $\varepsilon \in (0,1)$, dividing the inequality

$$\begin{aligned}
\int_{\Omega\times\mathbb{S}^{d-1}} & G(x,v,(1-\varepsilon)m^{\overline{Q}}(x,v) + \varepsilon m^Q(x,v))\,\theta(dx,dv) \\
&\qquad - \int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m^{\overline{Q}}(x,v))\,\theta(dx,dv) \geq 0
\end{aligned}$$

by ε and letting $\varepsilon \to 0^+$ we obtain

$$\int_{\Omega\times\mathbb{S}^{d-1}} \xi_{\overline{Q}} \cdot m^{\overline{Q}} = \int_{\mathcal{L}} L_{\xi_{\overline{Q}}} d\overline{Q} \leq \int_{\Omega\times\mathbb{S}^{d-1}} \xi_{\overline{Q}} \cdot m^Q = \int_{\mathcal{L}} L_{\xi_{\overline{Q}}} dQ,\ \forall Q \in \mathcal{Q}^q(\gamma).$$

Adapting the proof of Proposition 3.9 of [39] to our case, we see that the infimum of the right-hand side of the previous inequality is in fact $\int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\xi_{\overline{Q}}} d\gamma$ so that we have

$$\int_{\mathcal{L}} L_{\xi_{\overline{Q}}} d\overline{Q} = \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0),\sigma(1))\,d\overline{Q}(\sigma,\rho)$$

3. Let $\overline{Q}$ solve (3.13). Then considering $\xi = \xi_{\overline{Q}}$ and $m = m^{\overline{Q}}$, inequality (3.11) becomes an equality and (3.10) as well due to (3.14). Thus (3.12) is in fact an equality and $\xi_{\overline{Q}}$ solves (3.1). □

A natural question is to investigate the discrete problems corresponding to (2.4) i.e.

$$\inf_{\boldsymbol{m}^\varepsilon,\boldsymbol{w}^\varepsilon} \sum_{(x,e)\in E^\varepsilon} |e|^d G\left(x, \frac{e}{|e|}, \frac{m^\varepsilon(x,e)}{|e|^{d/2}}\right) \tag{3.22}$$

subject to the mass conservation conditions (2.2)-(2.3) and convergence of problems (3.22) in some sense to the continuous problem

$$\inf_{Q\in\mathcal{Q}(\gamma)} \int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m^Q(x,v))\,\theta(dx,dv). \tag{3.23}$$

Let $\mathbf{m}^\varepsilon = (m^\varepsilon(x,e))_{(x,e)\in E^\varepsilon}$ and $\mathbf{w}^\varepsilon = (w^\varepsilon(\sigma))_{\sigma\in C^\varepsilon}$ solve the discrete problem (3.22). Let $\sigma = (x_0,\dots,x_{L(\sigma)}) \in C^\varepsilon$ (identified with the piecewise affine curve defined on $[0, L(\sigma)]$). For every $k = 0,\dots,L(\sigma)-1$, let us denote by i_k the integer such that $(x_k, x_{k+1} - x_k) \in E^\varepsilon_{i_k}$. Then let us define $\rho^\sigma \in L^\infty([0,1])^N$ where for all $t \in [k, k+1[$,

$$\rho^\sigma_i(t) = \begin{cases} |\sigma(k+1) - \sigma(k)| & \text{if } i = i_k, \\ 0 & \text{otherwise.} \end{cases}$$

We will define a discrete measure Q^ε over $\mathcal{L}^\varepsilon$ where

$$\mathcal{L}^\varepsilon := \{(\sigma, \rho^\sigma) : \sigma \in C^\varepsilon\}.$$

Write Q^ε as follows

$$Q^\varepsilon := \varepsilon^{d/2-1} \sum_{\sigma \in C^\varepsilon} w^\varepsilon(\sigma) \delta_{\sigma \otimes \rho^\sigma}$$

as well as

$$\widetilde{Q}^\varepsilon := \varepsilon^{d/2-1} \sum_{\sigma \in C^\varepsilon} w^\varepsilon(\sigma) \delta_{\tilde{\sigma} \otimes \rho^{\tilde{\sigma}}}$$

where $\tilde{\sigma} \in W^{1,\infty}([0,1], \overline{\Omega})$ is the constant speed reparameterization of the path σ. Notice that for every $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, we have $L_\xi(\sigma, \rho^\sigma) = L_\xi(\tilde{\sigma}, \rho^{\tilde{\sigma}})$ so that $m^{Q^\varepsilon} = m^{\widetilde{Q}^\varepsilon}$. Let us also observe that the measure $m^{\widetilde{Q}^\varepsilon}$ contains all the information on $(\mathbf{m}^\varepsilon, \mathbf{w}^\varepsilon)$.

Especially for the following theorem, we make a stronger assumption.

Assumption 3.1. *There exists a function $C : \mathbb{R}_+ \mapsto \mathbb{R}_+^*$ such that $C(\varepsilon) \to 1$ as $\varepsilon \to 0^+$ and for every $\varepsilon > 0, (x, e) \in E^\varepsilon, C(\varepsilon)\varepsilon \leq |e| \leq \varepsilon$.*

In particular, this hypothesis is satisfied in our three classical examples since arc length is constant for $\varepsilon > 0$ fixed.

Theorem 3.2. *Under the previous assumptions, defining $\widetilde{Q}^\varepsilon$ as above, up to a subsequence, $(\widetilde{Q}^\varepsilon)_{\varepsilon>0}$ weakly converges to some solution $Q \in \mathcal{Q}^q(\gamma)$ of (3.23) in the sense that*

$$\int_{C([0,1],\mathbb{R}^d)\times L^1([0,1])^N} \Phi(\sigma,\rho) d\widetilde{Q}^\varepsilon(\sigma,\rho) \to \int_{C([0,1],\mathbb{R}^d)\times L^1([0,1])^N} \Phi(\sigma,\rho) dQ(\sigma,\rho),$$

as $\varepsilon \to 0^+$ for every $\Phi \in C_b(C([0,1], \mathbb{R}^d) \times L^1([0,1])^N, \mathbb{R})$.

Proof. By duality, from Theorem 3.1 and Corollary 2.1, it follows that the value of (3.22) converges to that of (3.23) and in particular, due to the q growth condition (3.5) on $G(x, v, \cdot)$, $\mathbf{m}^\varepsilon$ is bounded for the discrete L^q norm. In the same manner that in the proof of Corollary 2.1 and Section 4.1 in Chapter 2 we can see that there is some $m \in L^q_+$ such that $(x, e) \to \frac{m^\varepsilon(x,e)}{|e|^{d/2}}$ weakly converges to m in L^q in the sense of definition 2.2 (up to replacing p by q) and

$$\int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m(x,v))\,\theta(dx,dv) \leq \liminf_{\varepsilon\to 0^+} \sum_{(x,e)\in E^\varepsilon} |e|^d G\left(x, \frac{e}{|e|}, \frac{m^\varepsilon(x,e)}{|e|^{d/2}}\right). \quad (3.24)$$

Let $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R}_+)$, recalling (2.3), (2.10) and (4.9), rearranging terms, we have

$$\begin{aligned}
&\int_{\Omega\times\mathbb{S}^{d-1}} \xi(x,v)\, dm^{\widetilde{Q}^\varepsilon}(x,v) = \int_{\mathcal{L}} L_\xi(\sigma,\rho)\, d\widetilde{Q}^\varepsilon(\sigma,\rho)\\
&= \varepsilon^{d/2-1} \sum_{\sigma\in C^\varepsilon} w^\varepsilon(\sigma) \sum_{k=0}^{L(\sigma)-1} \int_k^{k+1} \xi(\sigma(t), v_{i_k}(\sigma(t)))|\sigma(k+1)-\sigma(k)|dt\\
&= \varepsilon^{d/2-1} \sum_{\sigma\in C^\varepsilon} w^\varepsilon(\sigma) \sum_{k=0}^{L(\sigma)-1} \int_{[\sigma(k),\sigma(k+1)]} \xi\left(\cdot, \frac{\sigma(k+1)-\sigma(k)}{|\sigma(k+1)-\sigma(k)|}\right) + O(\omega_\xi(\varepsilon))\\
&= \varepsilon^{d/2-1} \sum_{\sigma\in C^\varepsilon} w^\varepsilon(\sigma) \sum_{k=0}^{L(\sigma)-1} \left(\xi\left(\sigma(k), \frac{\sigma(k+1)-\sigma(k)}{|\sigma(k+1)-\sigma(k)|}\right) + O(\omega_\xi(\varepsilon))\right) |\sigma(k+1)-\sigma(k)|\\
&= \varepsilon^{d/2-1} \sum_{(x,e)\in E^\varepsilon} \left(\xi\left(x, \frac{e}{|e|}\right) + O(\omega_\xi(\varepsilon))\right) \left(\sum_{\sigma\in C^\varepsilon:(x,e)\subset\sigma} |e| w^\varepsilon(\sigma)\right)\\
&= \varepsilon^{d/2-1} \sum_{(x,e)\in E^\varepsilon} |e|^{d/2+1} \xi\left(x, \frac{e}{|e|}\right) \frac{m^\varepsilon(x,e)}{|e|^{d/2}} + O(\omega_\xi(\varepsilon))
\end{aligned}$$

where ω_ξ is a modulus of continuity of ξ. From Assumption 3.1 and the fact that $(x,e) \to \frac{m^\varepsilon(x,e)}{|e|^{d/2}}$ weakly converges in L^q to m in the sense of definition 2.2, it follows that $m^{\widetilde{Q}^\varepsilon}$ weakly star converges to m. Arguing as previously, we find $Q \in \mathcal{M}_1^+(\mathcal{L})$ such that, up to a subsequence, $(\widetilde{Q}^\varepsilon)_\varepsilon$ weakly converges to Q and $m^Q \leq m$. We easily have $Q \in \mathcal{Q}^q(\gamma)$: indeed, for every $\varphi \in C(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$, we have

$$\begin{aligned}
\int_{\mathcal{L}} \varphi(\sigma(0),\sigma(1))\, dQ(\sigma,\rho) &= \lim_{\varepsilon\to 0^+} \int_{\mathcal{L}} \varphi(\sigma(0),\sigma(1))\, d\tilde{Q}^\varepsilon(\sigma,\rho)\\
&= \lim_{\varepsilon\to 0^+} \varepsilon^{d/2-1} \sum_{\sigma\in C^\varepsilon} w^\varepsilon(\sigma)\varphi(\sigma(0),\sigma(1))\\
&= \lim_{\varepsilon\to 0^+} \varepsilon^{d/2-1} \sum_{(x,y)\in {N^\varepsilon}^2} \varphi(x,y) \left(\sum_{\sigma\in C^\varepsilon_{x,y}} w^\varepsilon(\sigma)\right)\\
&= \lim_{\varepsilon\to 0^+} \varepsilon^{d/2-1} \sum_{(x,y)\in {N^\varepsilon}^2} \varphi(x,y)\gamma^\varepsilon(x,y)\\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \varphi\, d\gamma.
\end{aligned}$$

Using (3.24) and the fact that $G(x,v,\cdot)$ is nondecreasing, we get

$$\begin{aligned}
\int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m^Q(x,v))\,\theta(dx,dv) &\leq \int_{\Omega\times\mathbb{S}^{d-1}} G(x,v,m(x,v))\,\theta(dx,dv)\\
&\leq \liminf_{\varepsilon\to 0^+} \sum_{(x,e)\in E^\varepsilon} |e|^d G\left(x, \frac{e}{|e|}, \frac{m^\varepsilon(x,e)}{|e|^{d/2}}\right).
\end{aligned}$$

Since the right-hand side is the value of the infimum in (3.23), we obtain the desired result. □

2 The long-term variant

Instead of taking the transport plan γ^ε as given in the discrete problem, we now consider the case where only its marginals are fixed. More precisely, there is a distribution of sources $f_-^\varepsilon = \varepsilon^{d/2-1}\sum_{x\in N^\varepsilon} f_-^\varepsilon(x)\delta_x$ and sinks $f_+^\varepsilon = \varepsilon^{d/2-1}\sum_{x\in N^\varepsilon} f_+^\varepsilon(x)\delta_x$

which are discrete measures with same total mass on the set of nodes N^ε (that we can assume to be 1 as a normalization)

$$\sum_{x\in N^\varepsilon} f_-^\varepsilon(x) = \sum_{y\in N^\varepsilon} f_+^\varepsilon(y) = \varepsilon^{1-d/2}.$$

The numbers $f_-^\varepsilon(x)$ and $f_+^\varepsilon(x)$ are nonnegative for every $x \in N^\varepsilon$.

With the same notations as in the short-term problem, we have almost the same definition of an equilibrium as in definition 4.1, we must change the mass conservation condition (2.2) as follows

$$f_-^\varepsilon(x) := \sum_{\sigma\in C_{x,\cdot}^\varepsilon} w^\varepsilon(\sigma),\ f_+^\varepsilon(y) := \sum_{\sigma\in C_{\cdot,y}^\varepsilon} w^\varepsilon(\sigma) \tag{3.25}$$

for every $(x, y) \in N^\varepsilon \times N^\varepsilon$, where $C_{x,\cdot}^\varepsilon$ (respectively $C_{\cdot,y}^\varepsilon$) is the set of loop-free paths starting at the origin x (respectively stopping at the terminal point y). Moreover, the transport plan now is an unknown. Similar arguments apply to this case, the equilibrium is a minimizer of the functional defined by (2.4) but now subject to (3.25) and (2.3). We shall then state the analogue of the dual formulation (2.5)

$$\inf_{t^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}} \left\{ \sum_{(x,e)\in E^\varepsilon} H^\varepsilon(x, e, t^\varepsilon(x, e)) - \inf_{\gamma^\varepsilon\in\Pi(f_-^\varepsilon, f_+^\varepsilon)} \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x, y) T_{t^\varepsilon}^\varepsilon(x, y) \right\}, \tag{3.26}$$

where $\Pi(f_-^\varepsilon, f_+^\varepsilon)$ is the set of discrete transport plans between f_-^ε and f_+^ε, that is, the set of nonnegative numbers $(\gamma^\varepsilon(x, y))_{(x,y)\in N^{\varepsilon 2}}$ such that

$$\sum_{y\in N^\varepsilon} \gamma^\varepsilon(x, y) = f_-^\varepsilon(x),\ \sum_{x\in N^\varepsilon} \gamma^\varepsilon(x, y) = f_+^\varepsilon(y),\ \forall (x, y) \in N^\varepsilon \times N^\varepsilon,$$

We assume that the hypotheses made in Subsection 3.1 in Chapter 2 are still satisfied, except that we replace Assumption 2.8 by

Assumption 3.2. *f_-^ε and f_+^ε weakly star converge to some probability measures f_- and f_+ on $\overline{\Omega}$:*

$$\lim_{\varepsilon\to 0^+} \varepsilon^{d/2-1} \sum_{x\in N^\varepsilon} (\varphi(x) f_-^\varepsilon(x) + \psi(x) f_+^\varepsilon(x)) = \int_{\overline{\Omega}} \varphi df_- + \int_{\overline{\Omega}} \psi df_+,\ \forall (\varphi, \psi) \in C(\overline{\Omega})^2.$$

Writing ξ^ε as in (2.14), we can now reformulate (3.26)

$$\inf_{\xi^\varepsilon\in\mathbb{R}_+^{\#E^\varepsilon}} F^\varepsilon(\xi^\varepsilon) := I_0^\varepsilon(\xi^\varepsilon) - F_1^\varepsilon(\xi^\varepsilon) \tag{3.27}$$

where $I_0^\varepsilon(\xi^\varepsilon)$ is defined by (2.19) and

$$F_1^\varepsilon(\xi^\varepsilon) := \inf_{\gamma^\varepsilon\in\Pi(f_-^\varepsilon, f_+^\varepsilon)} \sum_{(x,y)\in N^{\varepsilon 2}} \gamma^\varepsilon(x, y) \left(\min_{\sigma\in C_{x,y}^\varepsilon} \sum_{(z,e)\subset\sigma} |e|^{d/2} \xi^\varepsilon(z, e) \right). \tag{3.28}$$

It is an optimal transport problem. The limit functional then reads as the following variant of (2.25)

$$F(\xi) := I_0(\xi) - F_1(\xi),\ \text{where } F_1(\xi) := \inf_{\gamma\in\Pi(f_-, f_+)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi d\gamma,\ \forall \xi \in L_+^p(\theta), \tag{3.29}$$

As previously, I_0 is defined by (2.21) and $\overline{c}_\xi$ by (2.24). $\Pi(f_-, f_+)$ is the set of transport plans between f_- and f_+ (see (3.16)). We then have the following Γ-convergence result :

Theorem 3.3. *Under the same assumptions except Assumption 2.8 replaced by Assumption 3.2, the family of functionals F^ε defined by* (3.27) Γ*-converges (for the weak L^p-topology) to the functional F defined by* (3.29).

Proof. We can do the same reasoning as for Theorem 2.2 except for the proof of the inequality

$$F_1(\xi) \geq \limsup_{\varepsilon} F_1^\varepsilon(\xi^\varepsilon)$$

when $\xi^\varepsilon \to \xi$. In order to show it, we use the following lemma whose proof is in [12].

Lemma 3.7. *Let μ and ν be probability measures on $\overline{\Omega}$, $(\mu_n)_n$, $(\nu_n)_n$ be sequences of probability measures on $\overline{\Omega}$ that weakly-$\star$ converge to μ and ν and let $\gamma \in \Pi(\mu, \nu)$ a transport plan between μ and ν. Then there exists a sequence of transport plans $\gamma_n \in \Pi(\mu_n, \nu_n)$ that weakly star converges to γ.*

In much the same way as in Subsection 4.1 in Chapter 2, c^ε given by (2.32) has a subsequence that converges uniformly to some $c \leq \overline{c}_\xi$. Then let us take $\gamma \in \Pi(f_-, f_+)$ such that

$$F_1(\xi) = \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi d\gamma.$$

It follows from Lemma 3.7 that there exists a sequence $(\gamma^\varepsilon) \in \Pi(f_-^\varepsilon, f_+^\varepsilon)$ such that γ^ε weakly-$\star$ converges to γ as $\varepsilon \to 0$. We then have

$$\begin{aligned}\limsup_{\varepsilon} F_1^\varepsilon(\xi^\varepsilon) &\leq \limsup_{\varepsilon} \sum_{(x,y)\in N^\varepsilon \times N^\varepsilon} \gamma^\varepsilon(x,y) c^\varepsilon(x,y) \\ &= \int_{\overline{\Omega}\times\overline{\Omega}} c\, d\gamma \leq \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_\xi d\gamma = F_1(\xi).\end{aligned}$$

□

In the same manner as in Section 1 of Chapter 2, we can see that the problem (3.29) has a dual formulation that is

$$\sup_{Q\in\mathcal{Q}^q(f_-,f_+)} -\int_{\Omega\times\mathbb{S}^{d-1}} G(x, v, m^Q(x, v))\theta(dx, dv), \tag{3.30}$$

where

$$\begin{aligned}\mathcal{Q}^q(f_-, f_+) &:= \{Q \in \mathcal{M}_+^1(\mathcal{L}) : e_{0\#}Q = f_-, e_{1\#}Q = f_+, m^Q \in L^q(\theta)\} \\ &= \bigcup_{\gamma\in\Pi(f_-,f_+)} \mathcal{Q}^q(\gamma).\end{aligned}$$

If we assume that $\mathcal{Q}^q(f_-, f_+) \neq \emptyset$ and that (3.5) is still true, one can reformulate Theorem 3.1 for the long-term models as follows :

Theorem 3.4. *We have :*

1. *Problem* (3.30) *admits solutions,*
2. $\overline{Q} \in \mathcal{Q}^q(f_-, f_+)$ *solves* (3.30) *if and only if*

$$\int_{\mathcal{L}} L_{\xi_{\overline{Q}}}(\sigma, \rho)\, d\overline{Q}(\sigma, \rho) = \int_{\mathcal{L}} \overline{c}_{\xi_{\overline{Q}}}(\sigma(0), \sigma(1))\, d\overline{Q}(\sigma, \rho)$$

where $\xi_{\overline{Q}}(x,v) := g(x,v,m^{\overline{Q}}(x,v))$ *and moreover,* $\overline{\gamma} := (e_0,e_1)_\# \overline{Q}$ *is a solution of the optimal transport problem:*

$$\inf_{\gamma\in\Pi(f_-,f_+)} \int_{\overline{\Omega}\times\overline{\Omega}} \overline{c}_{\xi_{\overline{Q}}}(x,y)d\gamma(x,y).$$

3. *There is no duality gap : the infimum of* (3.29) *equals the supremum of* (3.30) *and moreover, if* $\overline{Q}$ *solves* (3.30) *then* $\xi_{\overline{Q}}$ *solves* (3.29).

Problem (3.30) is studied in [66]. It is showed that problem (3.30) is equivalent to another problem that is the variational formulation of an anisotropic, degenerate and elliptic PDE :

$$\begin{cases} -\operatorname{div}\,(\nabla\mathcal{G}^*(x,\nabla u(x))) = f & \text{in } \Omega, \\ \nabla\mathcal{G}^*(x,\nabla u(x))\cdot\nu_\Omega = 0 & \text{on } \partial\Omega, \end{cases}$$

with $\mathcal{G}^*$ being a C^1 function. In particular, if the function g in (2.12) is of the form $g(x,v_k(x),m) = a_k(x)m^{q-1} + \delta_k$ for every $x\in\overline{\Omega}, k=1,\dots,N$ and $m\geq 0$ where the constants δ_k are positive and the weights a_k are regular and positive, then we have

$$\mathcal{G}^*(x,z) = \sum_{k=1}^N \frac{b_k(x)}{p}(z\cdot v_k(x) - \delta_k c_k(x))_+^p \text{ for every } x\in\overline{\Omega}, z\in\mathbb{R}^d$$

where $b_k = (a_k c_k)^{-\frac{1}{q-1}}$. This case is interesting since numerical simulations can be performed as shown in [66].

Chapter 4

Wardrop equilibria : long-term variant, degenerate anisotropic PDEs and numerical approximations

This chapter is from the paper [66].

Abstract : As shown in [67], under some structural assumptions, working on congested traffic problems in general and increasingly dense networks leads, at the limit by Γ-convergence, to continuous minimization problems posed on measures on generalized curves. Here we show the equivalence with another problem that is the variational formulation of an anisotropic, degenerate and elliptic PDE. For particular cases, we prove a Sobolev regularity result for the minimizers of the minimization problem despite the strong degeneracy and anisotropy of the Euler-Lagrange equation of the dual. We extend the analysis of [31] to the general case. Finally, we use the method presented in [21] to make numerical simulations.
Keywords: traffic congestion, Wardrop equilibrium, generalized curves, anisotropic and degenerate PDEs, augmented Lagrangian.

1 Introduction

Researchers in the field of modeling traffic have developed the concept of congestion in networks since the early 50's and the introduction of the notion of Wardrop equilibrium (see [98]). Its wide popularity is due to some applications to road traffic and communication networks. We will describe the general congested network model built in [67] in the following subsection.

1.1 Presentation of the general discrete model

Given $d \in \mathbb{N}, d \geq 2$ and Ω a bounded domain of $\mathbb{R}^d$ with a Lipschitz boundary and $\varepsilon > 0$, we take a sequence of finite oriented networks $\Omega_\varepsilon = (N^\varepsilon, E^\varepsilon)$. The set of nodes in Ω_ε is N^ε and E^ε is the set of pairs (x, e) with $x \in N^\varepsilon$ and $e \in \mathbb{R}^d$ such that $|e|$ is of order ε and the segment $[x, x+e]$ is included in Ω. We will simply identify arcs to pairs (x, e). We assume $|E^\varepsilon| := \max\{|e|, \text{ there exists } x \text{ such that } (x, e) \in E^\varepsilon\} = \varepsilon$.

Masses and congestion: Let us denote the traffic flow on the arc (x, e) by $m^\varepsilon(x, e)$. There is a function $g^\varepsilon : E^\varepsilon \times \mathbb{R}_+ \to \mathbb{R}_+$ such that for each $(x, e) \in E^\varepsilon$ and $m \geq 0$, $g^\varepsilon(x, e, m)$ represents the traveling time of arc (x, e) when the mass on (x, e) is m. The function g^ε is positive and increasing in its last variable. This describes the congestion effect. We will denote the collection of all arc-masses $m^\varepsilon(x, e)$ by $\mathbf{m}^\varepsilon$.

Marginals: There is a distribution of sources $f_\varepsilon^- = \sum_{x \in N^\varepsilon} f_\varepsilon^-(x)\delta_x$ and sinks $f_\varepsilon^+ = \sum_{x \in N^\varepsilon} f_\varepsilon^+(x)\delta_x$ which are discrete measures with same total mass on the set of nodes N^ε (that we can assume to be 1 as a normalization)

$$\sum_{x \in N^\varepsilon} f_\varepsilon^-(x) = \sum_{y \in N^\varepsilon} f_\varepsilon^+(y) = 1.$$

The numbers $f_\varepsilon^-(x)$ and $f_\varepsilon^+(x)$ are nonnegative for every $x \in N^\varepsilon$.

Paths and equilibria: A path is a finite set of successive arcs $(x, e) \in E^\varepsilon$ on the network. C^ε is the finite set of loop-free paths on Ω_ε and may be partitioned as

$$C^\varepsilon = \bigcup_{(x,y) \in N^\varepsilon \times N^\varepsilon} C^\varepsilon_{x,y} = \bigcup_{x \in N^\varepsilon} C^\varepsilon_{x,\cdot} = \bigcup_{y \in N^\varepsilon} C^\varepsilon_{\cdot,y},$$

where $C^\varepsilon_{x,\cdot}$ (respectively $C^\varepsilon_{\cdot,y}$) is the set of loop-free paths starting at the origin x (respectively stopping at the terminal point y) and $C^\varepsilon_{x,y}$ is the intersection of $C^\varepsilon_{x,\cdot}$ and $C^\varepsilon_{\cdot,y}$. Then the travel time of a path $\gamma \in C^\varepsilon$ is given by:

$$\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\gamma) := \sum_{(x,e) \subset \gamma} g^\varepsilon(x, e, m^\varepsilon(x, e)).$$

The mass commuting on the path $\gamma \in C^\varepsilon$ will be denoted $w^\varepsilon(\gamma)$. The collection of all path-masses $w^\varepsilon(\gamma)$ will be denoted $\mathbf{w}^\varepsilon$. We may define an equilibrium that satisfies optimality requirements compatible with the distribution of sources and sinks and such that all paths used minimize the traveling time between their extremities, taking into account the congestion effects. In other words, we have to impose mass conservation conditions that relate arc-masses, path-masses and the data f_ε^- and f_ε^+:

$$f_\varepsilon^-(x) = \sum_{\gamma \in C^\varepsilon_{x,\cdot}} w^\varepsilon(\gamma),\ f_\varepsilon^+(y) = \sum_{\gamma \in C^\varepsilon_{\cdot,y}} w^\varepsilon(\gamma),\ \forall (x, y) \in N^\varepsilon \times N^\varepsilon \tag{4.1}$$

and

$$m^\varepsilon(x,e) = \sum_{\gamma \in C^\varepsilon : (x,e) \subset \gamma} w^\varepsilon(\gamma), \forall (x,e) \in E^\varepsilon. \tag{4.2}$$

We define $T^\varepsilon_{\mathbf{m}^\varepsilon}$ to be the minimal length functional, that is:

$$T^\varepsilon_{\mathbf{m}^\varepsilon}(x,y) := \min_{\gamma \in C^\varepsilon_{x,y}} \sum_{(x,e) \subset \gamma} g^\varepsilon(x,e,m^\varepsilon(x,e).$$

Let $\Pi(f^-_\varepsilon, f^+_\varepsilon)$ be the set of discrete transport plans between f^-_ε and f^+_ε, that is, the set of collection of nonnegative elements $(\varphi^\varepsilon(x,y))_{(x,y)\in N^{\varepsilon 2}}$ such that

$$\sum_{y \in N^\varepsilon} \varphi^\varepsilon(x,y) = f^-_\varepsilon(x) \text{ and } \sum_{x \in N^\varepsilon} \varphi^\varepsilon(x,y) = f^+_\varepsilon(y), \text{ for every } (x,y) \in N^\varepsilon \times N^\varepsilon.$$

This results in the concept of Wardrop equilibrium that is defined precisely as follows:

Definition 4.1. *A Wardrop equilibrium is a configuration of nonnegative arc-masses* $\mathbf{m}^\varepsilon : (x,e) \to (m^\varepsilon(x,e))$ *and of nonnegative path-masses* $\mathbf{w}^\varepsilon : \gamma \to w^\varepsilon(\gamma)$*, that satisfy the mass conservation conditions* (4.1) *and* (4.2) *and such that:*

1. *For every* $(x,y) \in N^\varepsilon \times N^\varepsilon$ *and every* $\gamma \in C^\varepsilon_{x,y}$*, if* $w^\varepsilon(\gamma) > 0$ *then*

$$\tau^\varepsilon_{\mathbf{m}^\varepsilon}(\gamma) = \min_{\gamma' \in C^\varepsilon_{x,y}} \tau^\varepsilon_{\mathbf{m}^\varepsilon}(\gamma'), \tag{4.3}$$

2. *If we define* $\Pi^\varepsilon(x,y) = \sum_{\gamma \in C^\varepsilon_{x,y}} w^\varepsilon(\gamma)$ *then* Π^ε *is a minimizer of*

$$\inf_{\varphi^\varepsilon \in \Pi(f^-_\varepsilon, f^+_\varepsilon)} \sum_{(x,y)\in N^\varepsilon \times N^\varepsilon} \varphi^\varepsilon(x,y) T^\varepsilon_{\mathbf{m}^\varepsilon}(x,y). \tag{4.4}$$

Condition (4.3) means that users behave rationally and always use shortest paths, taking in consideration congestion, that is, travel times increase with the flow. In [12, 67], the main discrete model studied is short-term, that is, the transport plan is prescribed. Here we work with a long-term variant as in [31, 33]. It means that we have fixed only the marginals (that are f^-_ε and f^+_ε). So the transport plan now is an unknown and must be determined by some additional optimality condition that is (4.4). Condition (4.4) requires that there is an optimal transport plan between the fixed marginals for the transport cost induced by the congested metric. So we also have an optimal transportation problem.

1.2 Assumptions and preliminary results

A few years after the work of Wardrop, Beckmann, McGuire and Winsten [17] observed that Wardrop equilibria coincide with the minimizers of a convex optimization problem:

Theorem 4.1. *A flow configuration* $(\mathbf{w}^\varepsilon, \mathbf{m}^\varepsilon)$ *is a Wardrop equilibrium if and only if it minimizes*

$$\sum_{(x,e)\in E^\varepsilon} G^\varepsilon(x,e,m^\varepsilon(x,e)) \text{ where } G^\varepsilon(x,e,m) := \int_0^m g^\varepsilon(x,e,\alpha)d\alpha \tag{4.5}$$

subject to nonnegativity constraints and the mass conservation conditions (4.1)-(4.2).

The problem (4.5) is interesting since it easily implies existence results and numerical schemes. However, it requires knowing the whole path flow configuration $\mathbf{w}^\varepsilon$ so that it may quickly be untractable for dense networks. However a similar issue was recently studied in [67]. Under structural assumptions, it is shown that we may pass to a continuous limit which will simplify the structure. Here, we will not see all these hypothesis, only the main ones. So we refer to [67] for more details. The only noticeable difference is that we take here $\beta = 1$ in Remark 1 of [67] for physical reasons.

Assumption 4.1. *The discrete measures $(f_\varepsilon^-)_{\varepsilon>0}$ and $(f_\varepsilon^-)_{\varepsilon>0}$ weakly star converge to some probability measures f^- and f^+ on $\overline{\Omega}$:*

$$\lim_{\varepsilon\to 0^+} \sum_{x\in N^\varepsilon} (\varphi(x) f_\varepsilon^-(x) + \psi(x) f_\varepsilon^+(x)) = \int_{\overline{\Omega}} \varphi df^- + \int_{\overline{\Omega}} \psi df^+,\ \forall (\varphi,\psi) \in C(\overline{\Omega})^2.$$

Assumption 4.2. *There exists $N \in \mathbb{N}$, $\{v_k\}_{k=1,\dots,N} \in C^1(\overline{\Omega}, \mathbb{S}^{d-1})^N$ and $\{c_k\}_{k=1,\dots,N} \in C^1(\overline{\Omega}, \mathbb{R}_+^*)^N$ such that E^ε weakly converges in the sense that*

$$\lim_{\varepsilon\to 0^+} \sum_{(x,e)\in E^\varepsilon} |e|^d \varphi\left(x, \frac{e}{|e|}\right) = \int_{\Omega\times\mathbb{S}^{d-1}} \varphi(x,v)\, \theta(dx,dv), \forall \varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}),$$

where $\theta \in \mathcal{M}_+(\Omega \times \mathbb{S}^{d-1})$ and θ is of the form

$$\theta(dx,dv) = \sum_{k=1}^N c_k(x) \delta_{v_k(x)} dx.$$

Moreover, there exists a constant $C > 0$ such that for every $(x,z,\xi) \in \overline{\Omega} \times \mathbb{S}^{d-1} \times \mathbb{R}_+^N$, there exists $\bar{Z} \in \mathbb{R}_+^N$ such that $|\bar{Z}| \le C$ and

$$\bar{Z}\cdot\xi = \min\left\{ Z\cdot\xi; Z = (z_1,\dots,z_N) \in \mathbb{R}_+^N \text{ and } \sum_{k=1}^N z_k v_k(x) = z \right\}. \tag{4.6}$$

The c_k's are the volume coefficients and the v_k's are the directions in the network. The measure θ depends on the discretization of Ω, i.e. the sequence $\{\Omega_\varepsilon\}_\varepsilon$. The last sub-assumption (4.6) allows us to keep some control on an optimal conical decomposition of all $z \in \mathbb{R}^d$ in the family of directions $\{v_k(x)\}_k$ for every $x \in \overline{\Omega}$. There always exists a conical decomposition of z in $\{v_k(x)\}$, not too large with respect to z. The next assumption focuses on the congestion functions g^ε.

Assumption 4.3. *g^ε is of the form*

$$g^\varepsilon(x,e,m) = |e| g\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d-1}}\right),\ \forall \varepsilon > 0, (x,e) \in E^\varepsilon, m \ge 0 \tag{4.7}$$

where $g : \Omega \times \mathbb{S}^{d-1} \times \mathbb{R}_+ \mapsto \mathbb{R}$ is a given continuous, nonnegative function that is increasing in its last variable.

In $\mathbb{R}^2$, it is very natural : the traveling time on an arc of length $|e|$ is of order $|e|$ and depends on the flow per unit of length $m/|e|$. In $\mathbb{R}^d$, it is a bit less natural.

The traveling time always is of order $|e|$ but now depends on $m/|e|^{d-1}$. We have also removed the ε-dependence on the g^ε. We then have

$$G^\varepsilon(x,e,m) = |e|^d G\left(x, \frac{e}{|e|}, \frac{m}{|e|^{d-1}}\right) \text{ where } G(x,v,m) := \int_0^m g(x,v,\alpha)d\alpha.$$

We also add assumptions on G:

Assumption 4.4. *There exists a closed neighborhood U of $\overline{\Omega}$ such that for $k = 1,\ldots,N$, v_k may be extended on U in a function C^1 (still denoted v_k). Moreover, each function $(x,m) \in U \times \mathbb{R}_+ \mapsto G(x, v_k(x), m)$ is Carathéodory, convex nondecreasing in its second argument with $G(x, v_k(x), 0) = 0$ a.e. $x \in U$ and there exists $1 < q < d/(d-1)$ and two constants $0 < \lambda \leq \Lambda$ such that for every $(x,m) \in U \times \mathbb{R}_+$ one has*

$$\lambda(m^q - 1) \leq G(x,v,m) \leq \Lambda(m^q + 1). \tag{4.8}$$

The q-growth is natural since we want to work in L^q in the continuous limit. The condition on q has a technical reason. It means that the conjugate exponent p of q is $> d$, which allows us to use Morrey's inequality in the proof of the convergence (see [67]). The extension on U will serve to use regularization by convolution and Moser's flow argument. Examples of models that satisfy these assumptions are regular decompositions. In two-dimensional networks, there exists three different regular decompositions: cartesian, triangular and hexagonal. In these models, the length of an arc in E^ε is ε. The c_k's and v_k's are constant. In the cartesian case, $N = 4$, $(v_1, v_2, v_3, v_4) := ((1,0),(0,1),(-1,0),(0,-1))$ and $c_k = 1$ for $k = 1,\ldots,4$. For more details, see [67].

Now, before presenting the continuous limit problem, let us set some notations.

Let us write the set of generalized curves

$$\mathcal{L} := \{(\gamma,\rho) : \gamma \in W^{1,\infty}([0,1],\overline{\Omega}), \rho \in \mathcal{P}_\gamma \cap L^\infty([0,1])^N\},$$

where

$$\mathcal{P}_\gamma := \left\{\rho : t \in [0,1] \to \rho(t) \in \mathbb{R}_+^N \text{ and } \dot{\gamma}(t) = \sum_{k=1}^N v_k(\gamma(t))\, \rho_k(t) \text{ a.e.}\right\}.$$

We can notice that $\mathcal{P}_\gamma$ is never empty thanks to Assumption 4.2. Let us denote $Q \in \mathcal{Q}(f^-, f^+)$ the set of Borel probability measures Q on $\mathcal{L}$ such that the mass conservation constraints are satisfied

$$\mathcal{Q}(f^-, f^+) := \{Q \in \mathcal{M}_+^1(\mathcal{L}) : e_{0\#}Q = f^-, e_{1\#}Q = f^+\}$$

where $e_t(\gamma,\rho) = \gamma(t), t \in [0,1], (\gamma,\rho) \in \mathcal{L}$. For $k = 1,\ldots,N$ let us then define the nonnegative measures on $\overline{\Omega} \times \mathbb{S}^{d-1}$, m_k^Q by

$$\int_{\overline{\Omega}\times\mathbb{S}^{d-1}} \varphi(x,v)dm_k^Q(x,v) := \int_{\mathcal{L}} \left(\int_0^1 \varphi(\gamma(t), v_k(\gamma(t)))\rho_k(t)dt\right) dQ(\gamma,\rho), \tag{4.9}$$

for every $\varphi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R})$. Then write simply $m^Q = \sum_{k=1}^N m_k^Q$, nonnegative measure on $\overline{\Omega} \times \mathbb{S}^{d-1}$. Finally assume that

$$\mathcal{Q}^q(f^-, f^+) := \{Q \in \mathcal{Q}(f^-, f^+) : m^Q \in L^q(\theta)\} \neq \emptyset.$$

It is true when for instance, f^+ and f^- are in $L^q(\Omega)$ and Ω is convex. Indeed, first for $Q \in \mathcal{M}^1_+(W^{1,\infty}([0,1],\overline{\Omega}))$, let us define $i_Q \in \mathcal{M}_+(\overline{\Omega})$ as follows

$$\int_\Omega \varphi \, di_Q := \int_{W^{1,\infty}([0,1],\overline{\Omega})} \left(\int_0^1 \varphi(\gamma(t))|\dot{\gamma}(t)|dt \right) dQ(\gamma) \text{ for } \varphi \in C(\overline{\Omega}, \mathbb{R}).$$

Thanks to the regularity results of [47, 89] there exists $Q \in \mathcal{M}^1_+(W^{1,\infty}([0,1],\overline{\Omega}))$ such that $e_{0\#}Q = f^-$, $e_{1\#}Q = f^+$ and $i_Q \in L^q$. For each curve γ, let $\rho^\gamma \in \mathcal{P}_\gamma$ such that $\sum_k \rho^\gamma_k(t) \leq C|\dot{\gamma}(t)|$ (we have the existence due to Assumption 4.2). Then we set $\tilde{Q} := (id, \rho^\cdot)_\# Q$. We have $\tilde{Q} \in \mathcal{Q}^q(f^-, f^+)$ so that we have proved the existence of such kind of measures. A necessary and sufficient condition to ensure $\mathcal{Q}^q(f^-, f^+) \neq \emptyset$ is that $f^+ - f^- \in W^{-1,q}(\Omega)$ (see [34]).

Then Wardrop equilibria at scale ε converge as $\varepsilon \to 0^+$ to solutions of the following problem

$$\inf_{Q \in \mathcal{Q}^q(f^-, f^+)} \int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, m^Q(x,v)) \theta(dx, dv) \tag{4.10}$$

(see [67]). Nevertheless this problem (4.10) is posed over probability measures on generalized curves and it is not obvious at all that it is simpler to solve than the discrete problem (4.5). So in the present paper, we want to show that problem (4.10) is equivalent to another problem that will roughly amount to solve an elliptic PDE. This problem is

$$\inf_{\sigma \in L^q(\Omega, \mathbb{R}^d)} \inf_{\varrho \in \mathcal{P}^\sigma} \left\{ \int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, \varrho(x,v)) \, \theta(dx, dv); \ -\mathrm{div} \ \sigma = f \right\}, \tag{4.11}$$

where

$$\mathcal{P}^\sigma := \left\{ \varrho : \Omega \times \mathbb{S}^{d-1} \to \mathbb{R}_+; \ \forall x \in \Omega, \ \sigma(x) = \sum_{k=1}^N v_k(x) \varrho(x, v_k(x)) \right\},$$

$f = f^+ - f^-$ and the equation $-\mathrm{div}(\sigma) = f$ is defined by duality:

$$\int_\Omega \nabla u \cdot \sigma = \int_\Omega u \, df, \text{ for all } u \in C^1(\overline{\Omega}),$$

so the homogeneous Neumann boundary condition $\sigma \cdot \nu_\Omega = 0$ is satisfied on $\partial\Omega$ in the weak sense. For the sake of clarity, let us define

$$\mathcal{G}(x, \sigma) := \inf_{\varrho \in \mathcal{P}^\sigma_x} \sum_{k=1}^N c_k(x) G(x, v_k(x), \varrho_k) := \inf_{\varrho \in \mathcal{P}^\sigma_x} \overline{G}(x, \varrho)$$

where

$$\mathcal{P}^\sigma_x := \left\{ \varrho \in \mathbb{R}^N_+; \ \sigma = \sum_{k=1}^N v_k(x) \varrho_k \right\} \text{ and } \overline{G}(x, \varrho) := \sum_{k=1}^N c_k(x) G(x, v_k(x), \varrho_k),$$

for $x \in \Omega, \sigma \in \mathbb{R}^d$. We recall that the c_k's are the volume coefficients in θ. $\mathcal{G}$ is convex in the second variable (since G is convex in its last variable).The minimization problem (4.11) can then be rewritten as

$$\inf_{\sigma \in L^q(\Omega, \mathbb{R}^d)} \left\{ \int_\Omega \mathcal{G}(x, \sigma(x)) \, dx; \ -\mathrm{div} \ \sigma = f \right\}. \tag{4.12}$$

This problem (4.12) looks like the ones introduced by Beckmann [16] for the design of an efficient commodity transport program. The dual problem of (4.12) takes the form

$$\sup_{u\in W^{1,p}(\Omega)} \left\{ \int_\Omega u \, df - \int_\Omega \mathcal{G}^*(x, \nabla u(x)) \, dx \right\}, \tag{4.13}$$

where p is the conjugate exponent of q and $\mathcal{G}^*$ is the Legendre transform of $\mathcal{G}(x,\cdot)$. In order to solve (4.12), we can first solve the Euler-Lagrange equation of its dual formulation and then use the primal-dual optimality conditions. Nevertheless, in our typical congestion models, the functions $G(x, v, \cdot)$ have a positive derivative at zero (that is $g(x, v, 0)$). Indeed, going at infinite speed - or teleportation - is not possible even when there is no congestion. So we have a singularity in the integrand in (4.12). Then G^* and the Euler-Lagrange equation of (4.13) are extremely degenerate. Moreover, the prototypical equation of [33] is the following

$$-\mathrm{div}\left((|\nabla u| - 1)_+^{p-1} \frac{\nabla u}{|\nabla u|} \right) = f.$$

Here, for well chosen g, we obtain anisotropic equation of the form

$$-\sum_{l=1}^{d} \partial_l \left[\sum_{k=1}^{N} (\nabla u \cdot v_k(x) - \delta_k c_k(x))_+^{p-1} v_k^l(x) \right] = f.$$

where $v_k(x) = (v_k^1(x), \dots, v_k^d(x))$ for $k = 1, \dots, N$ and $x \in \overline{\Omega}$. In the cartesian case (in $\mathbb{R}^2$), we can separate the variables in the sum (since here $\mathcal{G}(x, \sigma_1, \sigma_2) = \mathcal{G}_1(x, \sigma_1) + \mathcal{G}_2(x, \sigma_2)$). But in the hexagonal one ($d = 2$), it is impossible. The previous equation degenerates in an unbounded set of values of the gradient and its study is delicate, even if all the δ_k's are zero. It is more complicated than the one in [31]. Indeed, the studied model in [31] is the cartesian one and the prototypical equation is

$$-\sum_{k=1}^{2} \partial_k \left((|\partial_k u| - \delta_k)_+^{p-1} \frac{\partial_k u}{|\partial_k u|} \right) = f.$$

The plan of the paper is as follows. In Section 2, we formulate some relationship between (4.10) and (4.12). Section 3 is devoted to optimality conditions for (4.12) in terms of solutions of (4.13). We also present the kind of PDEs that represent realistic anisotropic models of congestion. In Section 4, we give some regularity results in the particular case where the c_k's and the v_k's are constant. Finally, in Section 5, we describe numerical schemes that allow us to approximate the solutions of the PDEs.

2 Equivalence with Beckmann problem

Let us study the relationship between problems (4.10) and (4.11). We still assume that all specified hypothesis in Section 1 are satisfied. Let us notice that thanks to Assumption 4.2, for every $\sigma \in L^q(\Omega, \mathbb{R}^d)$, there exists $\hat{\varrho} \in \mathcal{P}^\sigma$ such that $\hat{\varrho} \in L^q(\theta)$ and $\hat{\varrho}$ minimizes the following problem :

$$\inf_{\varrho\in\mathcal{P}^\sigma} \left\{ \int_{\Omega\times\mathbb{S}^{d-1}} G(x, v, \varrho(x, v)) \, \theta(dx, dv) \right\}.$$

For $\varrho \in \mathcal{P}^\sigma$, define $\bar{\varrho} : \Omega \to \mathbb{R}_+^N$ where $\bar{\varrho}_k(x) = \varrho(x, v_k(x))$, for every $x \in \Omega, k = 1, \dots, N$. Now, we only consider $\bar{\varrho}$ that we simply write ϱ (by abuse of notations).

Theorem 4.2. *Under all previous assumptions, we have*

$$\inf (4.10) = \inf (4.12).$$

Proof. We adapt the proof in [31]. We will show the two inequalities.
Step 1: $\inf (4.10) \geq \inf (4.12)$.
Let $Q \in \mathcal{Q}^q(f^-, f^+)$. We build $\sigma^Q \in L^q(\Omega, \mathbb{R}^d)$ that will allow us to obtain the desired inequality, we define it as follows :

$$\int_\Omega \varphi \, d\sigma^Q := \int_{\mathcal{L}} \int_0^1 \varphi(\gamma(t)) \cdot \dot{\gamma}(t) dt \, dQ(\gamma, \rho), \forall \varphi \in C(\overline{\Omega}, \mathbb{R}^d). \tag{4.14}$$

In particular, we have that $-\text{div } \sigma^Q = f$ since $Q \in \mathcal{Q}(f^-, f^+)$. We now justify that

$$\sigma^Q(x) = \int_{\mathbb{S}^{d-1}} v \, m^Q(x, v) \, dv = \sum_{k=1}^N v_k(x) m^Q(x, v_k(x)) \text{ a.e. } x \in \overline{\Omega}.$$

Recall that for every $\xi \in C(\overline{\Omega} \times \mathbb{S}^{d-1}, \mathbb{R})$,

$$\int_{\overline{\Omega} \times \mathbb{S}^{d-1}} \xi dm^Q = \int_{\mathcal{L}} \int_0^1 \left(\sum_{k=1}^N \xi(\gamma(t), v_k(\gamma(t))) \rho_k(t) \right) dt \, dQ(\gamma, \rho).$$

By taking ξ of the form $\xi(x, v) = \varphi(x) \cdot v$ with $\varphi \in C(\overline{\Omega}, \mathbb{R}^d)$, we get

$$\begin{aligned} \int_{\overline{\Omega} \times \mathbb{S}^{d-1}} \varphi(x) \cdot v \, dm^Q(x, v) &= \int_{\mathcal{L}} \int_0^1 \left(\sum_{k=1}^N \rho_k(t) \varphi(\gamma(t)) \cdot v_k(\gamma(t)) \right) dt \, dQ(\gamma, \rho) \\ &= \int_\Omega \varphi \, d\sigma^Q. \end{aligned}$$

Moreover, since $m^Q \geq 0$, we obtain that $m^Q \in \mathcal{P}^{\sigma^Q}$ (and so that $\sigma^Q \in L^q$) and the desired inequality follows.

Step 2: $\inf (4.10) \leq \inf (4.12)$.
Now prove the other inequality. We will use Moser's flow method (see [33, 44, 82]) and a classical regularization argument. Fix $\delta > 0$. Let $\sigma \in L^q(\Omega, \mathbb{R}^d)$ and $\varrho \in \mathcal{P}^\sigma \cap L^q(\Omega, \mathbb{R}^N)$ such that

$$\int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, \varrho(x, v)) \, \theta(dx, dv) \leq \inf \, (4.12) + \delta$$

with $-\text{div } \sigma = f$. We extend them outside Ω by 0. Let then $\eta \in C_c^\infty(\mathbb{R}^d)$ be a positive function, supported in the unit ball B_1 and such that $\int_{\mathbb{R}^d} \eta = 1$. For $\varepsilon \ll 1$ so that $\Omega_\varepsilon := \Omega + \varepsilon B_1 \Subset U$, we define $\eta^\varepsilon(x) := \varepsilon^{-d} \eta(\varepsilon^{-1} x)$, $\sigma^\varepsilon := \eta^\varepsilon \star \sigma$ and $\varrho_k^\varepsilon(x) := \eta^\varepsilon \star \varrho_k(x)$ for $k = 1, \dots, N$. By construction, we thus have that $\sigma^\varepsilon \in C^\infty(\overline{\Omega}_\varepsilon)$ and

$$- \text{ div } (\sigma^\varepsilon) = f_\varepsilon^+ - f_\varepsilon^- \text{ in } \Omega_\varepsilon \text{ and } \sigma^\varepsilon = 0 \text{ on } \partial\Omega_\varepsilon,$$

where $f^\varepsilon_\pm = \eta^\varepsilon \star (f_\pm 1_{\overline{\Omega}}) + \varepsilon$. But the problem is that we do not have $\varrho^\varepsilon \in \mathcal{P}^{\sigma^\varepsilon}$. We shall build a sequence (P^ε) in $\mathcal{P}^{\sigma^\varepsilon}$ that converges to ρ in $L^q(U, \mathbb{R}^N)$. Notice that

$$\begin{aligned}\sigma^\varepsilon(x) &= \sum_{k=1}^N \int \eta^\varepsilon(y)\varrho_k(x-y)v_k(x-y)\,dy \\ &= \sum_{k=1}^N \varrho^\varepsilon_k(x)v_k(x) + \sum_{k=1}^N \int \eta^\varepsilon(y)\varrho_k(x-y)(v_k(x-y)-v_k(x))\,dy\end{aligned}$$

There exists $p^\varepsilon_k \in L^q(\Omega_\varepsilon)$ such that for every $k = 1, \dots, N$, $p^\varepsilon_k \geq 0, p^\varepsilon_k \to 0$ and for $x \in \Omega_\varepsilon$, we have

$$I^\varepsilon(x) = \sum_{k=1}^N \int \eta^\varepsilon(y)\varrho_k(x-y)(v_k(x-y)-v_k(x))\,dy = \sum_{k=1}^N p^\varepsilon_k(x)v_k(x).$$

Such a family exists since $I^\varepsilon \in L^q$ and $I^\varepsilon \to 0$ (by using the fact that the v_k's are in $C^1(U)$) and we can estimate p^ε_k with I^ε due to Assumption 4.2. Then if we set $P^\varepsilon = \varrho^\varepsilon + p^\varepsilon$, we have $P^\varepsilon \in \mathcal{P}^{\sigma^\varepsilon}$ and $P^\varepsilon \to \varrho$ in L^q.

Define $g^\varepsilon(t,x) := (1-t)f^-_\varepsilon(x) + tf^+_\varepsilon(x)\,\forall t \in [0,1]$, $x \in \overline{\Omega}_\varepsilon$, let then X^ε be the flow of the vector field $v^\varepsilon := \sigma^\varepsilon/g^\varepsilon$, that is,

$$\begin{cases}\dot{X}^\varepsilon_t(x) = v^\varepsilon(t, X^\varepsilon_t(x)) \\ X^\varepsilon_0(x) = x, \;\; (t,x) \in [0,1] \times \overline{\Omega}_\varepsilon.\end{cases}$$

We have $\partial_t g^\varepsilon + \mathrm{div}\,(g^\varepsilon v^\varepsilon) = 0$. Since v^ε is smooth and the initial data is $g^\varepsilon(0,\cdot) = f^-_\varepsilon$, we have $X^\varepsilon_{t\,\#} f^-_\varepsilon = g^\varepsilon(t,\cdot)$. Let us define the set of generalized curves

$$\mathcal{L}_\varepsilon := \{(\gamma, \rho) : \gamma \in W^{1,\infty}([0,1], \Omega_\varepsilon), \rho \in \mathcal{P}_\gamma \cap L^\infty([0,1])^N\}.$$

Let us consider the following measure Q^ε on $\mathcal{L}_\varepsilon$

$$Q^\varepsilon := \int_{\overline{\Omega}_\varepsilon} \delta_{(X^\varepsilon_\cdot(x), P^\varepsilon(X^\varepsilon_\cdot(x))/g^\varepsilon(\cdot, X^\varepsilon_\cdot(x)))} df^-_\varepsilon(x).$$

We then have $e_{t\#}Q^\varepsilon = X^\varepsilon_{t\,\#} f^-_\varepsilon = g^\varepsilon(t,\cdot)$ for $t \in [0,1]$. We define σ^{Q^ε} and $m^{Q^\varepsilon}_k$ as in (4.14) and (4.9) respectively, by using test-functions defined on Ω_ε. We then have $\sigma^{Q^\varepsilon} = \sigma^\varepsilon$. Indeed, for $\varphi \in C(\overline{\Omega}_\varepsilon, \mathbb{R}^d)$, we have

$$\begin{aligned}\int_{\overline{\Omega}_\varepsilon} \varphi\, d\sigma^{Q^\varepsilon} &= \int_{\overline{\Omega}_\varepsilon} \int_0^1 \varphi(X^\varepsilon_t(x)) \cdot v^\varepsilon(t, X^\varepsilon_t(x)) f^-_\varepsilon(x)\,dt\,dx \\ &= \int_0^1 \int_{\overline{\Omega}_\varepsilon} \varphi(x) \cdot v^\varepsilon(t,x) g^\varepsilon(t,x)\,dx\,dt \\ &= \int_{\overline{\Omega}_\varepsilon} \varphi\, d\sigma^\varepsilon\end{aligned}$$

which gives the equality. We used the definition of Q^ε, the fact that $X^\varepsilon_{t\,\#} f^-_\varepsilon = g^\varepsilon(t,\cdot)$ and that $v^\varepsilon g^\varepsilon = \sigma^\varepsilon$ and Fubini's theorem. In the same way, we have $m^{Q^\varepsilon} \in \mathcal{P}^{\sigma^\varepsilon}$. To prove it, we take the same arguments as in the end of Step 1 and in the previous calculation. For $\varphi \in C(\Omega_\varepsilon, \mathbb{R}^d)$, we have

$$\begin{aligned}
&\int_{\Omega_\varepsilon \times \mathbb{S}^{d-1}} \varphi(x) \cdot v \, m^{Q^\varepsilon}(dx, dv) \\
&\qquad = \int_0^1 \left(\int_{\Omega_\varepsilon} \sum_{k=1}^N \varphi(X_t^\varepsilon(x)) \cdot v_k(X_t^\varepsilon(x)) \frac{P_k^\varepsilon(X_t^\varepsilon(x))}{g^\varepsilon(t, X_t^\varepsilon(x))} f_\varepsilon^-(x) dx \right) dt \\
&\qquad = \int_0^1 \left(\int_{\Omega_\varepsilon} \varphi(X_t^\varepsilon(x)) \cdot \frac{\sigma^\varepsilon(X_t^\varepsilon(x))}{g^\varepsilon(t, X_t^\varepsilon(x))} f_\varepsilon^-(x) dx \right) dt \\
&\qquad = \int_0^1 \left(\int_{\Omega_\varepsilon} \varphi(x) \cdot \sigma^\varepsilon(x) dx \right) dt \\
&\qquad = \int_{\Omega_\varepsilon} \varphi \, d\sigma^\varepsilon.
\end{aligned}$$

Moreover, more precisely, we have $m_k^{Q^\varepsilon}(dx, dv) = \delta_{v_k(x)} P_k^\varepsilon(x) dx$. Then we conclude as in [31]. First for any Lipschitz curve φ, let us denote by $\tilde{\varphi}$ its constant speed reparameterization, that is, for $t \in [0,1], \tilde{\varphi}(t) = \varphi(s^{-1}(t))$, where

$$s(t) = \frac{1}{l(\varphi)} \int_0^t |\dot{\varphi}(u)| du \text{ with } l(\varphi) = \int_0^1 |\dot{\varphi}(u)| du.$$

For $(\varphi, \rho) \in \mathcal{L}$, let $\tilde{\rho}$ be the reparameterization of ρ i.e.

$$\tilde{\rho}_k(t) := \frac{l(\sigma)}{|\dot{\sigma}(s^{-1}(t))|} \rho_k(s^{-1}(t)), \forall t \in [0,1], k = 1, \ldots, N.$$

Let us denote by $\tilde{Q}$ the push forward of Q through the map $(\varphi, \rho) \mapsto (\tilde{\varphi}, \tilde{\rho})$. We have $m_k^{\tilde{Q}} = m_k^Q$ and $\sigma^{\tilde{Q}} = \sigma^Q$. Then arguing as in [67], the L^q bound on m^{Q^ε} yields the tightness of the family of Borel measures $\tilde{Q}^\varepsilon$ on $C([0,1], \mathbb{R}^d) \times L^\infty([0,1])^N$. So Q^ε $\star$-weakly converges to some measure Q (up to a subsequence). Let us remark that $\tilde{Q}^\varepsilon$ has its total mass equal to that of f_ε^+, that is, $1 + \varepsilon|\Omega_\varepsilon|$. Thus one can show that $Q(\mathcal{L}) = 1$) (due to the fact that $Q(\mathcal{L}) = \lim_{\varepsilon \to 0^+} Q(\mathcal{L}_\varepsilon) = 1$). Moreover, we have $Q \in \mathcal{Q}(f^-, f^+)$ thanks to the $\star$-weak convergence of $\tilde{Q}^\varepsilon$ to Q. Recalling the fact that $P_k^\varepsilon = m^{Q^\varepsilon}(\cdot, v_k(\cdot))$ strongly converges in L^q to ϱ_k ($\varrho \in \mathcal{P}^\sigma$) and due to the same semicontinuity argument as in [39, 67], we have $m^Q(\cdot, v_k(\cdot)) \leq \varrho_k$ in the sense of measures. Then $m^Q(\cdot, v_k(\cdot)) \in L^q$ so that $Q \in \mathcal{Q}^q(f^-, f^+)$. It follows from the monotonicity of $G(x, v, \cdot)$ that :

$$\begin{aligned}
\int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, m^Q(x, v)) \, \theta(dx, dv) &\leq \int_{\Omega \times \mathbb{S}^{d-1}} G(x, v, \varrho(x, v)) \, \theta(dx, dv) \\
&\leq \inf (4.12) + \delta.
\end{aligned}$$

Letting $\delta \to 0^+$, we have the desired result. □

In fact, we showed in the previous proof a stronger result. We proved the following equivalence

$$Q \text{ solves } (4.10) \Longleftrightarrow \sigma^Q \text{ solves } (4.11)$$

and moreover,

$$(m^Q(\cdot, v_k(\cdot)))_{k=1,\ldots,N} \in \mathcal{P}^{\sigma^Q}$$

is optimal for (4.10). We also built a minimizing sequence for (4.10) from a regularization of a solution σ of (4.11) by using Moser's flow argument.

3 Characterization of minimizers via anisotropic elliptic PDEs

Here, we study the primal problem (4.12) and its dual problem (4.13). Recalling that $f = f^+ - f^-$ has zero mean, we can reduce the problem (4.13) only to zero-mean $W^{1,p}(\Omega)$ functions. Besides for $(x, v) \in \overline{\Omega} \times \mathbb{S}^{d-1}$ and $k = 1, \dots, N$, the functions $g_k(x, v, \cdot)$ are continuous positive and increasing on $\mathbb{R}_+$ since it is the time per unit of length to leave from the point x in the direction v when the intensity of traffic in this direction is m. Since $G(x, v, \cdot)$ has a positive derivative (that is $g_k(x, v, \cdot)$), G is strictly convex in its last variable then so is $\mathcal{G}(x, \cdot)$ for $x \in \overline{\Omega}$. Thus $\mathcal{G}^*$ is C^1. However $\mathcal{G}$ is not differentiable so that $\mathcal{G}^*(x, \cdot)$ is degenerate. By standard convex duality (Fenchel-Rockafellar's theorem, see [54] for instance), we have that $\min(4.12) = \max(4.13)$ and we can characterize the optimal solution σ of (4.12) (unique by strict convexity) as follows

$$\sigma(x) = \nabla \mathcal{G}^*(x, \nabla u(x)),$$

where u is a solution of (4.13). In other terms, u is a weak solution of the Euler-Lagrange equation

$$\begin{cases} -\operatorname{div}\left(\nabla \mathcal{G}^*(x, \nabla u(x))\right) = f & \text{in } \Omega, \\ \nabla \mathcal{G}^*(x, \nabla u(x)) \cdot \nu_\Omega = 0 & \text{on } \partial\Omega, \end{cases}$$

in the sense that

$$\int_\Omega \nabla \mathcal{G}^*(x, \nabla u(x)) \cdot \nabla \varphi(x)\, dx = \int_\Omega \varphi(x)\, df(x),\ \forall \varphi \in W^{1,p}(\Omega).$$

Let us remark that if u is not unique, σ is.

A typical example is $g(x, v_k(x), m) = g_k(x, m) = a_k(x) m^{q-1} + \delta_k$ with $\delta_k > 0$ and the weights a_k are regular and positive. We can explicitly compute $\mathcal{G}^*(x, z)$. Let us notice that for every $x \in \Omega, z \in \mathbb{R}^d$, we have :

$$\begin{aligned} \mathcal{G}^*(x, z) &= \sup_{\sigma \in \mathbb{R}^d} (z \cdot \sigma - \mathcal{G}(x, \sigma)) = \sup_{\sigma \in \mathbb{R}^d} (z \cdot \sigma - \inf_{\varrho \in \mathcal{P}_x^\sigma} \overline{G}(x, \varrho)) \\ &= \sup_{\sigma, \varrho} (z \cdot \sigma - \overline{G}(x, \varrho)) = \sup_{\varrho \in \mathbb{R}_+^N} \left\{ \sum_{k=1}^N (z \cdot v_k(x)) \varrho_k - \overline{G}(x, \varrho) \right\}. \end{aligned}$$

A direct calculus then gives

$$\mathcal{G}^*(x, z) = \sum_{k=1}^N \frac{b_k(x)}{p} (z \cdot v_k(x) - \delta_k c_k(x))_+^p,$$

where $b_k = (a_k c_k)^{-\frac{1}{q-1}}$. The PDE then becomes

$$-\sum_{k=1}^N \sum_{l=1}^d \partial_l \left[b_k(x) v_k^l(x) (\nabla u \cdot v_k(x) - \delta_k c_k(x))_+^{p-1} \right] = f, \tag{4.15}$$

where $v_k(x) = (v_k^1(x), \dots, v_k^d(x))$.

For $k = 1, \dots, N$, $\mathcal{G}_k^*(x,z) = \frac{b_k(x)}{p}(z \cdot v_k(x) - \delta_k)_+^p$ vanishes if $z \cdot v_k(x) \in]-\infty, \delta_k c_k(x)]$ so that any u whose the gradient satisfies $\nabla u(x) \cdot v_k(x) \in]-\infty, \delta_k c_k(x)]$, for all $x \in \Omega, k = 1, \dots, N$ is a solution of the previous PDE with $f = 0$. In consequence, we cannot hope to obtain estimates on the second derivatives of u or even oscillation estimates on ∇u from (4.15). Nevertheless we will see that we have some regularity results on the vector field $\sigma = (\sigma_1, \dots, \sigma_d)$ that solves (4.12) in the case where the directions and the volume coefficients are constant, that is,

$$\sigma(x) = \sum_{k=1}^{N} \left[b_k(x)(\nabla u(x) \cdot v_k - \delta_k c_k)_+^{p-1} \right] v_k,$$

for every $x \in \Omega$.

4 Regularity when the v_k's and c_k's are constant

Our aim here is to get some regularity results in the case where the v_k's and the c_k's are constant. We will strongly base on [31] to prove this regularity result. Let us consider the model equation

$$-\sum_{k=1}^{N} \operatorname{div}\left((\nabla u(x) \cdot v_k - \delta_k c_k)_+^{p-1} v_k \right) = f, \tag{4.16}$$

where $v_k \in \mathbb{S}^{d-1}, c_k > 0$ and $b_k \equiv 1$ for $k = 1, \dots, N$. Define for $z \in \mathbb{R}^d$

$$F(z) = \sum_{k=1}^{N} F_k(z), \text{ with } F_k(z) = (z \cdot v_k - \delta_k c_k)_+^{p-1} v_k \tag{4.17}$$

and

$$H(z) = \sum_{k=1}^{N} H_k(z), \text{ with } H_k(z) = (z \cdot v_k - \delta_k c_k)_+^{\frac{p}{2}} v_k. \tag{4.18}$$

Here we assume only $p \geq 2$. We have the following lemma that establishes some connections between F and H.

Lemma 4.1. *Let F and G be defined as above with $p \geq 2$, then for every $(z, w) \in \mathbb{R}^d \times \mathbb{R}^d$, the following inequalities are true for $k = 1, \dots, N$*

$$|F_k(z)| \leq |z|^{p-1}, \tag{4.19}$$

$$|F_k(z) - F_k(w)| \leq (p-1) \left(|H_k(z)|^{\frac{p-2}{p}} + |H_k(z)|^{\frac{p-2}{p}} \right) |H_k(z) - H_k(w)|, \tag{4.20}$$

and

$$(F_k(z) - F_k(w)) \cdot (z - w) \geq \frac{4}{p^2} |H_k(z) - H_k(w)|^2. \tag{4.21}$$

Proof. The first one is trivial. For the second one, from [76] one has the general result: for all $(a, b) \in \mathbb{R}^d \times \mathbb{R}^d$, the following inequality holds

$$\left| |a|^{p-2}a - |b|^{p-2}b \right| \leq (p-1) \left(|a|^{\frac{p-2}{2}} + |b|^{\frac{p-2}{2}} \right) \left| |a|^{\frac{p-2}{2}}a - |b|^{\frac{p-2}{2}}b \right|. \tag{4.22}$$

Choosing $a = (z \cdot v_k - \delta_k c_k)_+ v_k$ and $b = (w \cdot v_k - \delta_k c_k)_+ v_k$ in (4.22), we then obtain (4.20).

Let us now prove the third inequality. It is trivial if both $z \cdot v_k$ and $w \cdot v_k$ are less than $\delta_k c_k$. If $z \cdot v_k > \delta_k c_k$ and $w \cdot v_k \leq \delta_k c_k$, we have

$$(F_k(z) - F_k(w)) \cdot (z - w) = (z \cdot v_k - \delta_k c_k)_+^{p-1}(z \cdot v_k - w \cdot v_k) \geq (z \cdot v_k - \delta_k c_k)_+^p = |H_k(z)|^2.$$

For the case $z \cdot v_k > \delta_k c_k$ and $w \cdot v_k > \delta_k c_k$, we use the following inequality (again [76])

$$(|a|^{p-2}a - |b|^{p-2}b) \cdot (a - b) \geq \frac{4}{p^2}\left(|a|^{\frac{p-2}{2}}a - |b|^{\frac{p-2}{2}}b\right)^2.$$

Again taking $a = (z \cdot v_k - \delta_k c_k)_+ v_k$ and $b = (w \cdot v_k - \delta_k c_k)_+ v_k$, we have that

$$\begin{aligned}
&\frac{4}{p^2}|H_k(z) - H_k(w)|^2 \\
&\qquad \leq (|F_k(z)| - |F_k(w)|)v_k \cdot ((z \cdot v_k - \delta_k c_k)_+ - (w \cdot v_k - \delta_k c_k)_+)v_k \\
&\qquad = (|F_k(z)| - |F_k(w)|)(z - w) \cdot v_k,
\end{aligned}$$

which gives (4.21). □

Let us fix $f \in W^{1,q}_{\mathrm{loc}}(\Omega)$ where q is the conjugate exponent of p and let us consider the equation

$$-\operatorname{div} F(\nabla u) = f. \tag{4.23}$$

Thanks to Nirenberg's method of incremental ratios, we then have the following result that is strongly inspired of Theorem 4.1 in [31]:

Theorem 4.3. *Let $u \in W^{1,p}_{loc}(\Omega)$ be a local weak solution of* (4.23). *Then $\mathcal{H} := H(\nabla u) \in W^{1,2}_{loc}(\Omega)$. More precisely, for every $k = 1, \ldots, N, \mathcal{H}_k := H_k(\nabla u) \in W^{1,2}_{loc}(\Omega)$.*

Proof. For the sake of clarity, write $\mathcal{F} := F(\nabla u)$ and similarly, $\mathcal{F}_k, \mathcal{H}_k$ (note that $\mathcal{F}_k \in L^q_{\mathrm{loc}}(\Omega)$ and $\mathcal{H}_k \in L^2_{\mathrm{loc}}(\Omega)$ due to (4.19)-(4.20). Let us define the translate of the function φ by the vector h by $\tau_h\varphi := \varphi(\cdot + h)$. Let $\varphi \in W^{1,q}(\Omega)$ be compactly supported in Ω and $h \in \mathbb{R}^d \backslash \{0\}$ be such that $|h| < \mathrm{dist}(\mathrm{supp}(\varphi), \mathbb{R}^d \backslash \{0\})$, we then have

$$\int_\Omega \frac{\tau_h \mathcal{F} - \mathcal{F}}{|h|} \cdot \nabla\varphi dx = \int_\Omega \frac{\tau_h f - f}{|h|} \cdot \varphi dx. \tag{4.24}$$

Let $\omega \Subset \omega_0 \Subset \Omega$ and $\xi \in C^\infty_c(\Omega)$ such that $\mathrm{supp}(\xi) \subset \omega_0, 0 \leq \xi \leq 1$ and $\xi = 1$ on $\overline{\omega}$ and $h \in \mathbb{R}^d \backslash \{0\}$ such that $|h| \leq r_0 < \frac{1}{2}\mathrm{dist}(\omega_0, \mathbb{R}^d \backslash \Omega)$. In what follows, we denote by C a nonnegative constant that does not depend on h but may change from one line to another. We then introduce the test function

$$\varphi = \xi^2 |h|^{-1}(\tau_h u - u),$$

in (4.24). Let us fix $\omega' := \omega_0 + B(0, r_0)$. It follows from $u \in W^{1,p}_{\mathrm{loc}}(\Omega), f \in W^{1,q}_{\mathrm{loc}}(\Omega)$ and the Hölder inequality that

$$|h|^{-2}\int_\Omega (\tau_h \mathcal{F} - \mathcal{F}) \cdot \left(\xi^2(\tau_h \nabla u - \nabla u) + 2\xi\nabla\xi(\tau_h u - u)\right) \leq \|\nabla f\|_{L^q(\omega')}\|\nabla u\|_{L^p(\omega')}.$$

The left-hand side of the previous inequality is the sum of $2N$ terms $I_{11}+I_{12}+\ldots+I_{N1}+I_{N2}$ where for every $k=1,\ldots,N$,

$$I_{k1} := |h|^{-2}\int_\Omega \xi^2(F_k(\tau_h\nabla u)-F_k(\nabla u)\cdot(\tau_h\nabla u-\nabla u),$$

and

$$I_{k2} := |h|^{-2}\int_\Omega \xi^2(F_k(\tau_h\nabla u)-F_k(\nabla u)\cdot\nabla\xi\xi(\tau_h u-u).$$

Let $k=1,\ldots,N$ fixed. We will find estimations on I_{k1} and I_{k2}. Due to (4.20), I_{k1} satisfies:

$$I_{k1}\geq \frac{4}{p^2}\|\xi|h|^{-1}(\tau_h\mathcal{H}_k-\mathcal{H}_k)\|_{L^2}^2.$$

For I_{k2}, if $p>2$, it follows from (4.21) and the Hölder inequality with exponents $2,p$ and $2p/(p-2)$ that

$$\begin{aligned}
|I_{k2}| &\leq |h|^{-2}\int_\Omega |\xi\nabla\xi||\tau_h u-u||\tau_h\mathcal{H}_k-\mathcal{H}_k|\left(|\tau_h\mathcal{H}_k|^{\frac{p-2}{p}}+|\mathcal{H}_k|^{\frac{p-2}{p}}\right)\\
&\leq C\||h|^{-1}(\tau_h u-u)\|_{L^p(\omega_0)}\|\xi|h|^{-1}(\tau_h\mathcal{H}_k-\mathcal{H}_k)\|_{L^2}\left(\int_{\omega_0}|\mathcal{H}_k|^2+|\tau_h\mathcal{H}_k|^2\right)^{\frac{p-2}{2p}}\\
&\leq C\|\xi|h|^{-1}(\tau_h\mathcal{H}_k-\mathcal{H}_k)\|_{L^2},
\end{aligned}$$

and if $p=2$, we simply use Cauchy-Schwarz inequality and we get :

$$|I_{k2}|\leq C\|\xi|h|^{-1}(\tau_h\mathcal{H}_k-\mathcal{H}_k)\|_{L^2}.$$

Bringing together all estimates, we then obtain

$$\sum_{k=1}^N\left\|\xi\frac{\tau_h\mathcal{H}_k-\mathcal{H}_k}{h}\right\|_{L^2}^2\leq C\left(1+\sum_{k=1}^N\left\|\xi\frac{\tau_h\mathcal{H}_k-\mathcal{H}_k}{h}\right\|_{L^2}\right).$$

and we finally get

$$\sum_{k=1}^N\left\|\frac{\tau_h\mathcal{H}_k-\mathcal{H}_k}{h}\right\|_{L^2(\omega)}^2\leq C,$$

for some constant C that depends on p, $\|f\|_{W^{1,q}}$, $\|u\|_{W^{1,p}}$ and the distance between ω and $\partial\Omega$, but not on h. We have the desired result, that is, $\mathcal{H}_k\in W^{1,2}_{\rm loc}(\Omega)$, for $k=1,\ldots,N$, and so $\mathcal{H}$ also. $\square$

If we consider the variational problem of Beckmann type

$$\inf_{\sigma\in L^q(\Omega)}\left\{\int_\Omega\inf_{\varrho\in\mathcal{P}_x^\sigma}\sum_{k=1}^N c_k\left(\frac{1}{q}\varrho_k^q+\delta_k\varrho_k\right):-\mathrm{div}\,\sigma=f\right\},\tag{4.25}$$

we then have the following Sobolev regularity result for the unique minimizer that generalizes Corollary 4.3 in [31].

Corollary 4.1. *The solution σ of (4.25) is in the Sobolev space $W^{1,r}_{loc}(\Omega)$, where*

$$r=\begin{cases}2 & \text{if } p=2,\\ \text{any value } <2, & \text{if } p>2 \text{ and } d=2,\\ \dfrac{dp}{dp-(d+p)+2}, & \text{if } p>2 \text{ and } d>2.\end{cases}$$

Proof. By duality, we know the relation between σ and any solution of the dual problem u

$$\sigma = \sum_{k=1}^{N} (\nabla u \cdot v_k - \delta_k c_k)_+^{p-1} v_k.$$

Since $u \in W^{1,q}(\Omega)$ is a weak solution of the Euler-Lagrange equation (4.16), using Theorem 4.3 and Lemma 4.1, we have that the vector fields

$$\mathcal{H}_k(x) = (\nabla u(x) \cdot v_k - \delta_k c_k)_+^{\frac{p}{2}} v_k, \;\; k = 1, \dots, N,$$

are in $W^{1,2}_{\text{loc}}(\Omega)$. We then notice that $\sigma = \sum_{k=1}^{N} \sigma_k$ with

$$\sigma_k = |\mathcal{H}_k|^{\frac{p-2}{p}} \mathcal{H}_k, \;\; k = 1, \dots, N.$$

The first case is trivial: we simply have $\sigma_k = \mathcal{H}_k \in W^{1,2}_{\text{loc}}(\Omega)$. For the other cases, we use the Sobolev theorem. If $p > 2$ and $d > 2$ then $\mathcal{H}_k \in L^{2*}_{\text{loc}}(\Omega)$ with

$$\frac{1}{2^*} = \frac{1}{2} - \frac{1}{N}.$$

Applying (4.20) with $z = \tau_h \nabla u$ and $w = \nabla u$, we have

$$\left| \frac{\tau_h \sigma_k - \sigma_k}{|h|} \right| \leq (p-1) \left(|\tau_h \mathcal{H}_k|^{\frac{p-2}{p}} + |\mathcal{H}_k|^{\frac{p-2}{p}} \right) \left| \frac{\tau_h \mathcal{H}_k - \mathcal{H}_k}{|h|} \right|.$$

Since $|\mathcal{H}_k|^{\frac{p-2}{p}} \in L^{\frac{2^* p}{p-2}}_{\text{loc}}(\Omega)$, we have that the right-hand side term is in $L^r_{\text{loc}}(\Omega)$ with r given by

$$\frac{1}{r} = \frac{p-2}{2^* p} + \frac{1}{2}.$$

We can then control this integral

$$\int \left| \frac{\tau_h \sigma_k - \sigma_k}{|h|} \right|^r dx.$$

For the case $p > 2$ and $d = 2$, it follows from the same theorem that $\mathcal{H}_k \in L^s_{\text{loc}}(\Omega)$ for every $s < +\infty$ and the same reasoning allows us to conclude. $\square$

This Sobolev regularity result can be extended to equations with weights such as

$$-\sum_{k=1}^{N} \operatorname{div} \left(b_k(x) (\nabla u(x) \cdot v_k - \delta_k c_k)_+^{p-1} v_k \right) = f. \tag{4.26}$$

An open problem is to investigate if one can generalize this Sobolev regularity result to the case where the v_k's and c_k's are in $C^1(\overline{\Omega})$.

5 Numerical simulations

5.1 Description of the algorithm

We numerically approximate by finite elements solutions of the following minimization problem:

$$\inf_{u \in W^{1,p}(\Omega)} J(u) := \mathbf{G}^*(\nabla u) - \langle f, u \rangle \tag{4.27}$$

with $\mathbf{G}^*(\Phi) = \int_\Omega \mathcal{G}^*(x, \Phi(x))\, dx$ for $\Phi \in L^p(\Omega)^d$ and $\langle f, w\rangle = \int_\Omega u\, df$ for $w \in L^p(\Omega)$. Let us recall that Ω is a bounded domain of $\mathbb{R}^d$ with Lipschitz boundary and $f = f^+ - f^-$ is in the dual of $W^{1,p}(\Omega)$ with zero mean $\int_\Omega f = 0$. We will use the augmented Lagrangian method described in [21] (that we will recall later). ALG2 is a particular case of the Douglas-Rachford splitting method for the sum of two nonlinear operators (see [78] or more recently [84]). ALG2 was used for transport problems for the first time in [20]. Let a regular triangulation of Ω with typical meshsize h, let $E_h \subset W^{1,p}(\Omega)$ be the corresponding finite-dimensional space of P_2 finite elements of order 2 whose generic elements are denoted u_h. Moreover, if necessary, we approximate the terms f by $f_h \in E_h$ (again with $\langle f_h, 1\rangle = 0$) and $\mathbf{G}$ by a convex function $\mathbf{G}_h \in E_h$. Let us consider the approximating problem

$$\inf_{u_h \in E_h} J_h(u_h) := \mathbf{G}_h^*(\nabla u_h) - \langle f_h, u_h\rangle. \tag{4.28}$$

and its dual

$$\sup_{\sigma_h \in F_h^d} \{-\mathbf{G}_h(\sigma_h) : -\mathrm{div}_h(\sigma_h) = f_h\} \tag{4.29}$$

where F_h is the space of P_1 finite elements of order 1 and $-\mathrm{div}_h(\sigma_h)$ may be understood as

$$\langle \sigma_h, \nabla u_h\rangle_{F_h^d} = -\langle \mathrm{div}_h(\sigma_h), u_h\rangle_{E_h}.$$

Theorem 4.4. *If u_h solves* (4.28) *then up to a subsequence, u_h converges as $h \to 0$ to a u weakly in $W^{1,p}(\Omega)$ such that u solves* (4.27).

It is a direct application of a general theorem (see [21] and [61] for similar results and more details). Using the discretization by finite elements, (4.27) becomes

$$\inf_{u \in \mathbb{R}^n} J(u) := \mathbf{F}(u) + \mathbf{G}^*(\Lambda u) \tag{4.30}$$

where $\mathbf{F} : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \mathbf{G} : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ are two convex l.s.c. and proper functions and Λ is an $m \times n$ matrix with real entries. Λ is the discrete analogue of ∇. The dual of (4.30) then reads as

$$\sup_{\sigma \in \mathbb{R}^m} -\mathbf{F}^*(-\Lambda^T \sigma) - \mathbf{G}(\sigma) \tag{4.31}$$

We say that a pair $(\bar{u}, \bar{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m$ satisfies the primal-dual extremality relations if:

$$-\Lambda^T \bar{\sigma} \in \partial \mathbf{F}(\bar{u}), \bar{\sigma} \in \partial \mathbf{G}^*(\Lambda \bar{u}). \tag{4.32}$$

It means that $\bar{u}$ solves (4.30) and that $\bar{\sigma}$ solves (4.31) and moreover, (4.30) and (4.31) have the same value (no duality gap). It is equivalent to find a saddle-point of the augmented Lagrangian function for $r > 0$ (see [59, 61] for example)

$$L_r(u, q, \sigma) := \mathbf{F}(u) + \mathbf{G}^*(q) + \sigma \cdot (\Lambda u - q) + \frac{r}{2}|\Lambda u - q|^2,\ \forall (u, q, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m. \tag{4.33}$$

It is the discrete formulation of the corresponding augmented Lagrangian function

$$L_r(u, q, \sigma) := \int_\Omega \mathcal{G}^*(x, q(x))\, dx - \langle u, f\rangle + \langle \sigma, \nabla u - q\rangle + \frac{r}{2}|\nabla u - q|^2 \tag{4.34}$$

and the variational problem of (4.30) is

$$\inf_{u,q}\left\{\int_\Omega \mathcal{G}^*(x,q(x))\,dx - \int_\Omega u(x)f(x)\,dx\right\}. \tag{4.35}$$

subject to the constraint that $\nabla u = q$.

The augmented Lagrangian algorithm ALG2 involves building a sequence $(u^k, q^k, \sigma^k) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ from initial data (u^0, q^0, σ^0) as follows:

1. Minimization problem with respect to u:

$$u^{k+1} := \operatorname{argmin}_{u\in\mathbb{R}^n}\left\{\mathbf{F}(u) + \sigma^k\cdot\Lambda u + \frac{r}{2}|\Lambda u - q^k|^2)\right\}$$

   That is equivalent to solve the variational formulation of Laplace equation

$$-r(\Delta u^{k+1} - \operatorname{div}(q^k)) = f + \operatorname{div}(\sigma^k) \text{ in } \Omega$$

   with the Neumann boundary condition

$$r\frac{\partial u^{k+1}}{\partial \nu} = rq^k\cdot\nu - \sigma^k\cdot\nu \text{ on } \partial\Omega.$$

   This is where we use the Galerkin discretization by finite elements.

2. Minimization problem with respect to q:

$$q^{k+1} := \operatorname{argmin}_{q\in\mathbb{R}^m}\left\{\mathbf{G}^*(q) - \sigma^k\cdot q + \frac{r}{2}|\Lambda u^{k+1} - q|^2)\right\}$$

3. Using the gradient ascent formula for σ

$$\sigma^{k+1} = \sigma^k + r(\Lambda u^{k+1} - q^{k+1}).$$

Theorem 4.5. *Given $r > 0$. If there exists a solution to the primal-dual extremality relations* (4.32) *and Λ has full column-rank then there exists an $(\bar{u},\bar{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m$ satisfying* (4.32) *such that the sequence (u^k, q^k, σ^k) generated by the ALG2-scheme above satisfies*

$$u^k \to \bar{u}, q^k \to \Lambda\bar{u}, \sigma^k \to \bar{\sigma} \text{ as } k \to +\infty. \tag{4.36}$$

We directly apply a general theorem whose proof can be found in [53] (Theorem 8), following contributions of [59, 61, 78] to the analysis of splitting methods.

5.2 Numerical schemes and convergence study

We use the software FreeFem++ (see [68]) to implement the numerical scheme. We take the Lagrangian finite elements and notations used in Subsection 5.1, P_2 FE for u_h and P_1 FE for (q_h, σ_h). Λu_h is the projection on P_1 of the operator Λ, that is, ∇u_h. The first step and the third one are always the same and only the second one varies with our different test cases. We indicate the numerical convergence of ALG2 iterations by the $\cdot^k$ superscript and the convergence of finite elements discretization by the $\cdot_h$ subscript. For our numerical simulations, we work with the space dimension

$d = 2$ and we choose for Ω a $2D$ square ($x = (x_1, x_2) \in [0,1]^2$). We make tests with different f :

$$f_1^- := e^{-40*((x_1-0.75)^2+(x_2-0.25)^2} \text{ and } f_1^+ := e^{-40*((x_1-0.25)^2+(x_2-0.65)^2)},$$

$$f_2^- := e^{-40*((x_1-0.5)^2+(x_2-0.15)^2)} \text{ and } f_2^+ := e^{-40*((x_1-0.5)^2+(x_2-0.75)^2)},$$

In the third case, we take f_3^- a constant density and f_3^+ is the sum of three concentrated Gaussians

$$f_3^+(x_1,x_2) = e^{-400*((x-0.25)^2+(y-0.75)^2)} + e^{-400*((x-0.35)^2+(y-0.15)^2)} + e^{-400*((x-0.85)^2+(y-0.7)^2)}.$$

We also make tests with non-constant c_k :

$$h(x_1,x_2) = 3 - 2 * e^{-10*((x_1-0.5)^2+(y_2-0.5)^2)}.$$

As specified above, we use a triangulation of the unit square with $n = 1/h$ element on each side. We use the following convergence criteria:

1. DIV.Error $= \left(\int_{\Omega_h} (\mathrm{div}\sigma_h^k + f)^2\right)^{1/2}$ is the L^2 error on the divergence constraint.
2. BND.Error $= \left(\int_{\partial\Omega_h} (\sigma_h^k \cdot \nu)^2\right)^{1/2}$ is the $L^2(\partial\Omega_h)$ error on the Neumann boundary condition.
3. DUAL.Error $= \max_{x_j} |\mathcal{G}(x_j, \sigma_h^k(x_j)) + \mathcal{G}^*(x_j, \nabla u_h^k(x_j)) - \nabla u_h^k(x_j)\cdot\sigma_h^k(x_j)|$ where the maximum is with respect to the vertices x_j.

The first two criteria represent the optimality conditions for the minimization of the Lagrangian with respect to u and the third one is for maximization with respect to σ.

We make tests for two models. In the first one, the directions are the same as in the cartesian model and the volume coefficients are not necessarily constant. In the second one, the directions are the same than in the hexagonal one and the volume coefficients are equal to 1 (it is simpler to compute $\mathcal{G}(x,\sigma)$). That is, $v_k = \exp(ik\pi/3)$ and $\delta_k c_k = 1$ for $k = 1, \dots, 6$. We call these models still the cartesian one, the hexagonal one respectively. The cartesian one is much easier since we can separate variables. $\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2$ with $\mathbf{G}_i(x,q) = \frac{b_i}{p}(|q_i| - \delta_i c_i(x))_+^p$ so that the second step of ALG2 is equivalent to solve the pointwise problem

$$\inf_q \frac{1}{p}(|q| - c(x))_+^p + \frac{r}{2}|q - \tilde{q}^k|^2$$

where $\tilde{q}^k = \Lambda u^{k+1} + \frac{\sigma^k}{r}$. This amounts to set $q^{k+1} = \lambda\tilde{q}^k$ and to solve this equation in λ

$$(\lambda|\tilde{q}^k| - c(x))_+^{p-1} + r\lambda|\tilde{q}^k| = r|\tilde{q}^k| = 0$$

with $\lambda \geq 0$. We can use the dichotomy algorithm.

For the hexagonal one, we use Newton's method. Since the function of which we seek the minimizer has its Hessian matrix that is definite positive, we can use the inverse of this Hessian matrix.

Test case	DIV.Error	BND.Error	DUAL.Error	Time execution (seconds)
1	8.4745e-05	0	3.6126e-06	436
2	2.2536e-05	8.8705e-04	3.0663e-05	4764
3	5.2141e-05	1.4736e-04	1.1556e-02	792
4	1.1823e-05	7.6776e-04	8.7412e-06	170
5	1.1629e-05	0	9.7498e-04	285
6	3.5553e-04	1.2406	2.1083e-06	431
7	4.1373e-04	1.1710	4.8113e-04	4657

Table 4.1 – Convergence of the finite element discretization for all test cases.

We show the results of numerical simulations after 200 iterations for both models. All figures represent σ except the bottom image in Figure 4.1 and Figure 4.2 that shows q.

We notice that length of arrows are proportional to transport density. Level curves correspond to the density term of the source/sink data to be transported. In Figure 4.3, the case $p = 1.01$ means that there is much congestion. The case $p = 2$ is reasonable congestion and in the last one $p = 100$, there is little congestion. When there are obstacles, the criteria BND.Error is not very good. Indeed, the flow comes right on the obstacle and it turns fast. In the other side of the obstacle, the flow is tangent to the border. Many other cases may of course be examined (other boundary conditions, obstacles, coefficients depending on x, different exponents p for the different components of the flow...).

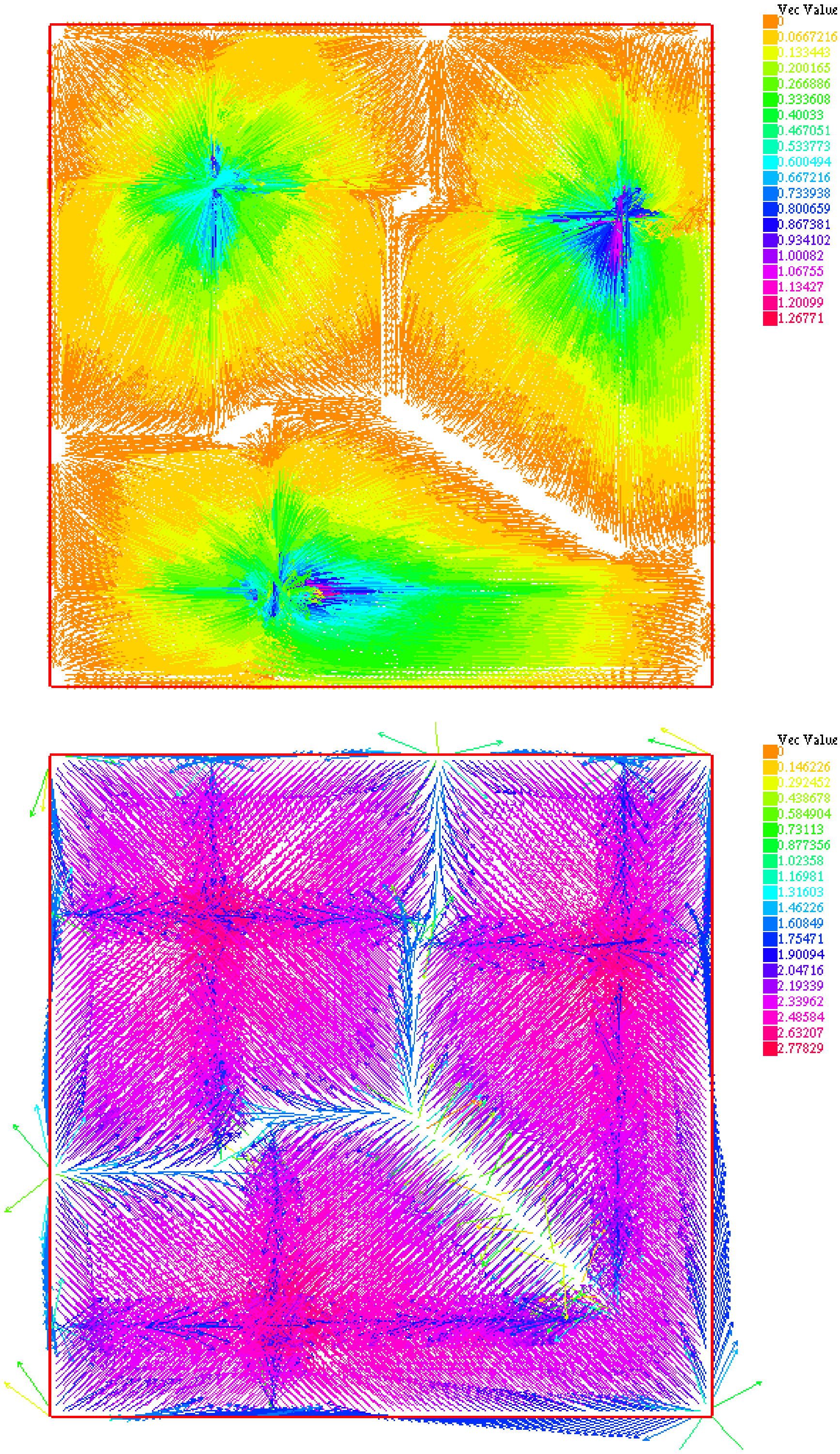

Figure 4.1 – Test case 1 : cartesian case ($d = 2$) with $f = f_3$, c_k constant and $p = 10$.

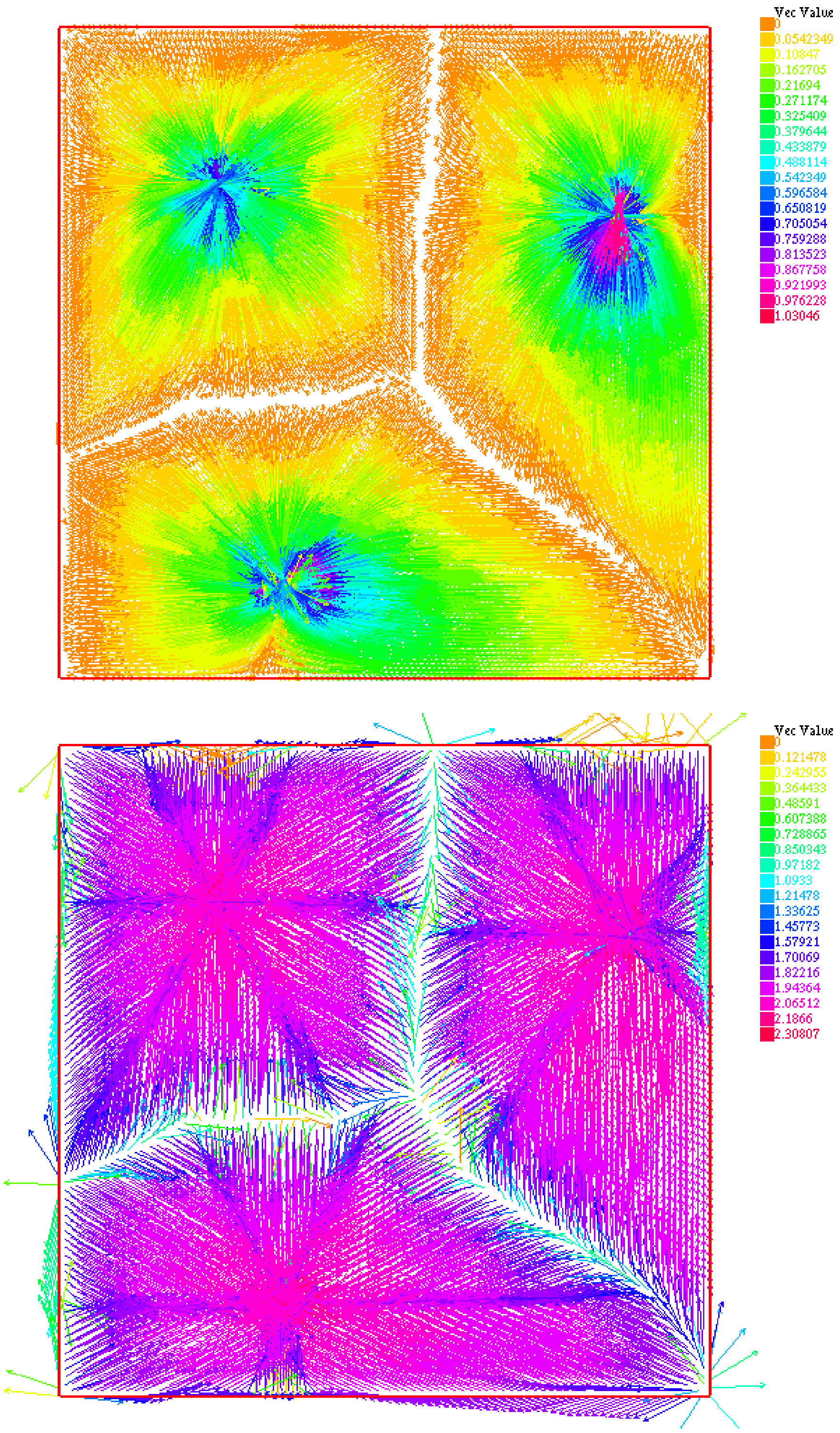

Figure 4.2 – Test case 2 : hexagonal case ($d = 2$) with $f = f_3$, c_k constant and $p = 3$.

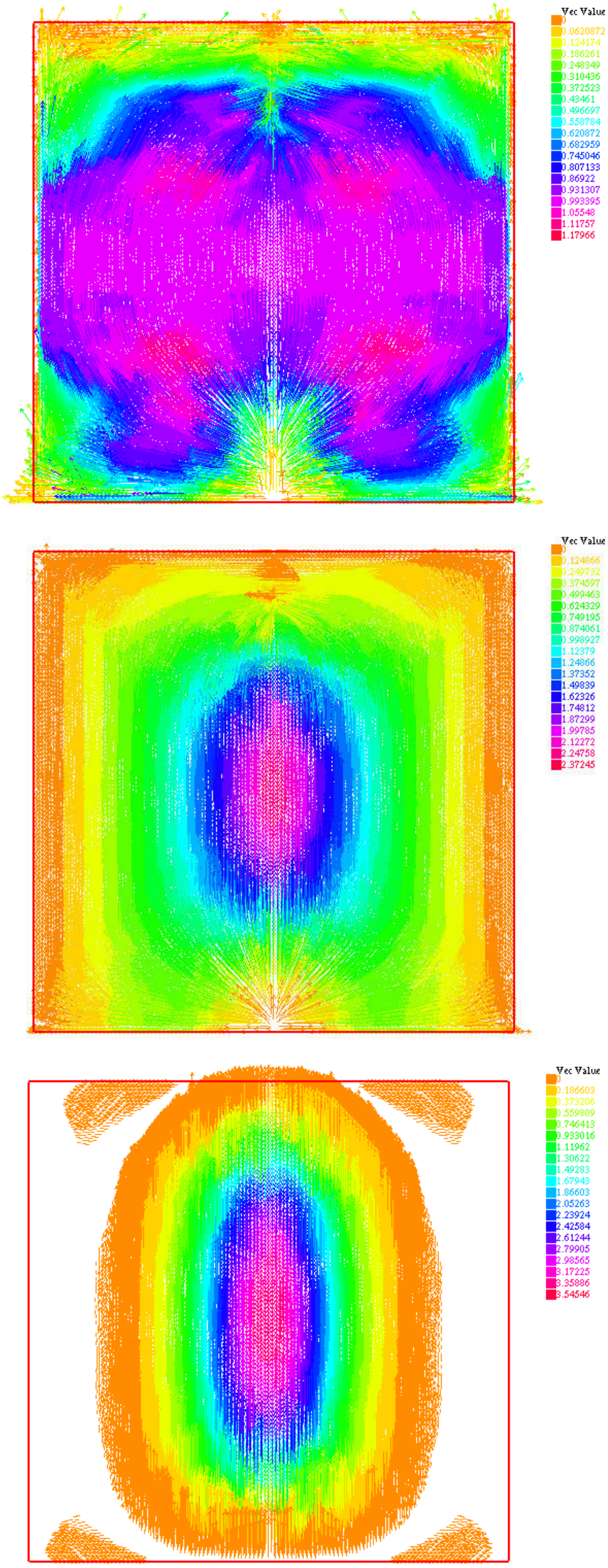

Figure 4.3 – Test cases 3, 4 and 5: cartesian case ($d = 2$) with $f = f_2$, c_k constant and $p = 1.01, 2, 100$.

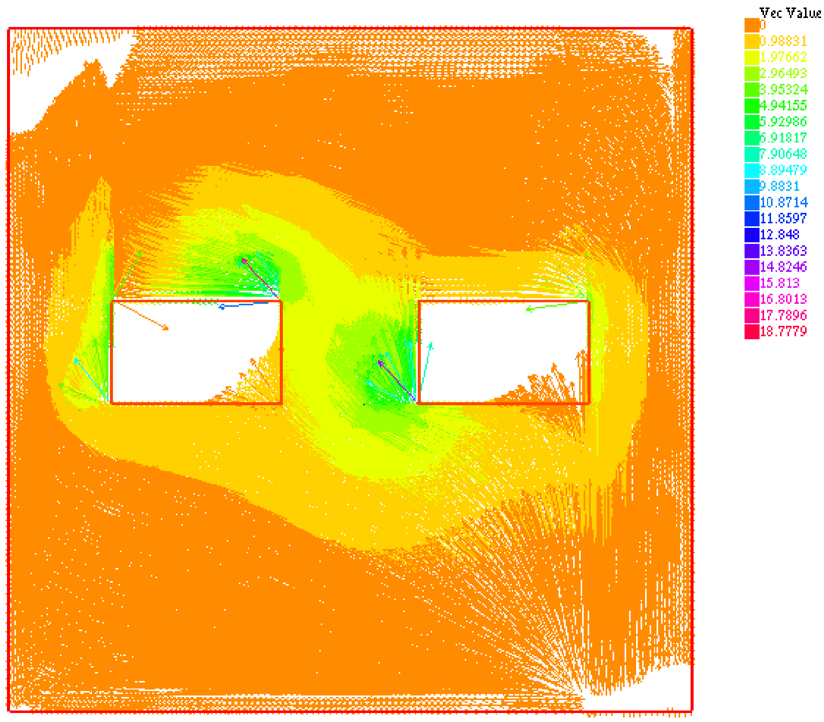

Figure 4.4 – Test case 6 : cartesian case ($d = 2$) with $f = f_1$, $c_1 = h$ and $c_2 = 1$, $p = 3$ and two obstacles.

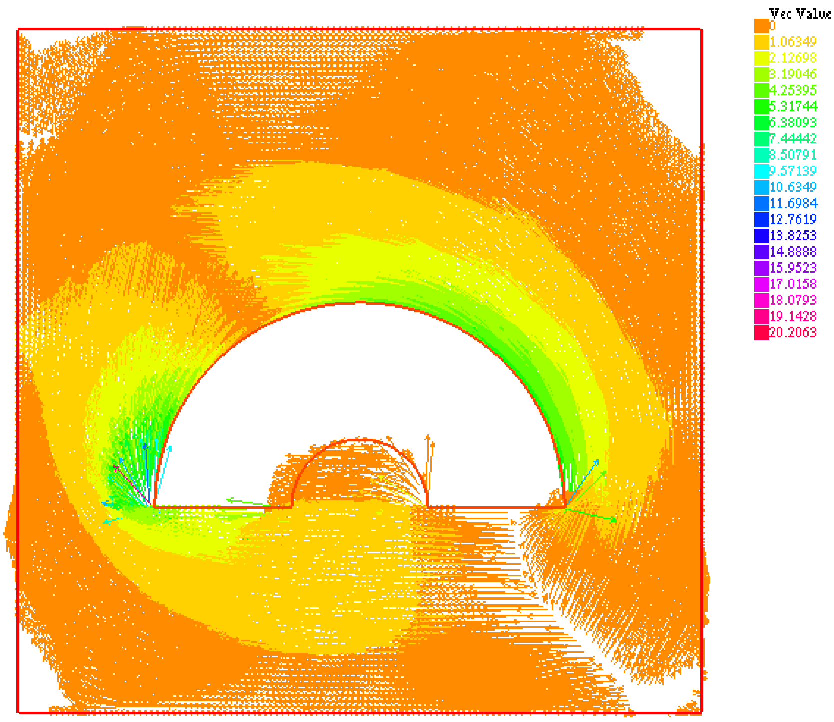

Figure 4.5 – Test case 7 : hexagonal case ($d = 2$) with $f = f_1$, c_k constant, $p = 3$ and an obstacle.

Chapter 5

A numerical solution to Monge's problem with Finsler distance as cost

This chapter is from the paper [22] written with Benamou and Carlier.

Abstract : Monge's problem with a Finsler cost is intimately related to an optimal flow problem. Discretization of this problem and its dual leads to a well-posed finite-dimensional saddle-point problem which can be solved numerically relatively easily by an augmented Lagrangian approach in the same spirit as the Benamou-Brenier method for the optimal transport problem with quadratic cost. Numerical results validate the method. We also emphasize that the algorithm only requires elementary operations and in particular never involves evaluation of the Finsler distance or of geodesics.

Keywords: Monge's problem, Finsler distance, augmented Lagrangian.

MS Classification: 65K10, 90C25, 90C46.

1 Introduction

Given a bounded domain Ω of $\mathbb{R}^d$, and two probability measures f^+ and f^- on $\overline{\Omega}$, we are interested in solving Monge's problem

$$\inf_{\pi\in\Pi(f^-,f^+)} \int_{\overline{\Omega}\times\overline{\Omega}} d_L(x,y)\mathrm{d}\pi(x,y) \tag{5.1}$$

where $\Pi(f^-, f^+)$ is the set of transport plans between f^- and f^+ i.e. the set of probability measures having f^- and f^+ as marginals and d_L is a Finsler distance. More precisely, d_L is given by

$$d_L(x,y) := \inf\left\{ \int_0^1 L(\gamma(t),\dot{\gamma}(t))\ \mathrm{d}t\ :\ \gamma\in W^{1,1}([0,1],\overline{\Omega}),\ \gamma(0)=x,\ \gamma(1)=y\right\} \tag{5.2}$$

where the Lagrangian L: $\overline{\Omega}\times\mathbb{R}^d\to\mathbb{R}_+$ is a continuous function of Finsler type i.e. for every $x\in\overline{\Omega}$, $v\mapsto L(x,v)$ is a norm and there is a constant $C>0$ such that the following nondegeneracy condition holds:

$$\frac{|v|}{C}\le L(x,v)\le C|v|,\ \forall(x,v)\in\overline{\Omega}\times\mathbb{R}^d. \tag{5.3}$$

Of course, one difficulty is the evaluation of the cost, and we shall see how to avoid computing it. This will be done by considering suitable dual, minimal flow and saddle-point formulations for which one can easily use an augmented Lagrangian method. The use of augmented Lagrangian methods in optimal transport was pioneered in the seminal work of Benamou and Brenier [20] on the dynamic formulation of the quadratic optimal transport case. For a distance cost (Monge case), in fact there is no need to introduce an additional time-variable and the analogue of the Benamou-Brenier dynamic problem is the minimal flow problem introduced by Beckmann [16]. We refer to the recent work [21] of the first two authors for applications of these augmented Lagrangian methods to Mean-Field-Games and optimal transport and to the work of the third author [66] for applications to anisotropic congested optimal transport. To the best of our knowledge, the relevance of augmented Lagrangian methods for a general Finsler metric has remained unnoticed in the literature. For other methods to solve optimal problems with the euclidean distance as transport cost, we refer for instance to [15] where a certain regularization is considered. Our goal is to show that Monge's problem with a Finsler metric is in fact quite easy to solve directly numerically by using an augmented Lagrangian approach.

The paper is organized as follows. In Section 2, we recall several reformulations of the Monge problem with Finsler cost (5.1): the Kantorovich dual, the minimal flow reformulation and finally a (formal) saddle-point problem for finding at the same time the Kantorovich potential and the optimal flow field. Section 3 describes the discretization saddle-point problem (which is well-posed), discusses the convergence and details the steps of the augmented algorithm ALG2 of Glowinski and Fortin. Section 4 gives numerical results.

2 Reformulations

2.1 Dual and minimal flow formulations

The standard Kantorovich duality formula (see [95]) says that the infimum in Monge's problem (5.1) coincides with the value of the dual:

$$\sup\left\{\langle u,f\rangle:=\int_{\overline{\Omega}}u(x)\mathrm{d}(f^+-f^-)(x)\ :\ u \text{ is 1-Lipschitz for } d_L\right\}. \tag{5.4}$$

Thanks to (5.3), it is easy to see that if u is 1-Lipschitz for d_L it is actually Lipschitz hence differentiable a.e., moreover the constraint $u(x)-u(y)\le d_L(x,y)$ can be expessed in differential form as follows. Defining the dual norm $L^*(x,.)$ of $L(x,.)$:

$$L^*(x,p):=\sup\{p\cdot v\ :\ L(x,v)\le 1\},$$

one can express the fact that u is 1-Lipschitz for d_L by the following pointwise constraint on ∇u

$$L^*(x,\nabla u(x))\le 1 \text{ for a.e. } x\in\overline{\Omega} \tag{5.5}$$

i.e.

$$\sigma\cdot\nabla u(x)\le L(x,\sigma),\ \forall\sigma\in\mathbb{R}^d.$$

Thus (5.4) can be rewritten in sup-inf form as

$$\sup_{u\in W^{1,\infty}}\inf_{\sigma\in L^1(\Omega,\mathbb{R}^d)}\langle u,f\rangle+\int_\Omega L(x,\sigma(x))\mathrm{d}x-\int_\Omega\nabla u(x)\cdot\sigma(x)\mathrm{d}x. \tag{5.6}$$

Switching the infimum and the supremum above, we obtain another dual formulation of (5.4):

$$\inf_{\sigma\in L^1(\Omega,\mathbb{R}^d)}\int_\Omega L(x,\sigma(x))\mathrm{d}x+\sup_{u\in W^{1,\infty}}\langle u,f\rangle-\int_\Omega\nabla u(x)\cdot\sigma(x)\mathrm{d}x$$

observing that the supremum with respect to u is 0 if $-\mathrm{div}(\sigma)=f=f^+-f^-$ and $\sigma\cdot\nu=0$ on $\partial\Omega$ in the weak sense i.e.

$$\int_{\overline{\Omega}}\nabla u(x)\cdot\sigma(x)\mathrm{d}x=\langle u,f\rangle,\ \forall u\in C^1(\overline{\Omega})$$

and $+\infty$ otherwise, we obtain the following minimal flow problem dual formulation of (5.4):

$$\inf_{\sigma\in L^1(\Omega,\mathbb{R}^d)}\left\{\int_\Omega L(x,\sigma(x))\mathrm{d}x\ :\ -\mathrm{div}(\sigma)=f\right\}. \tag{5.7}$$

Minimal flow formulations for transport problems were first introduced in the 1950's by Beckmann in an economic context [16], the connection with Monge's problem was realized much later by Robert McCann (see in particular [58]). It is obvious that (5.4) possesses solutions. Standard convex duality also implies that there is no duality gap and that

$$\sup(5.4)=\inf(5.1)=\inf(5.7). \tag{5.8}$$

It is however not clear in general that (5.7) possesses L^1 solutions. In the spatially homogeneous case where $L(x,v)$ is the euclidean norm (or more generally some smooth and uniformly convex norm), $|\sigma|$ is called the transport density and there are important and involved L^1 regularity results for the transport density under suitable assumptions on $f^{\pm}$ due to Feldman and McCann [58], De Pascale, Evans and Pratelli [47], De Pascale and Pratelli [49] and Santambrogio [89]. We are not aware of extensions to the Finsler case yet. Since the cost in (5.7) is convex and homogeneous of degree one, (5.7) can be relaxed to vector-valued measures which amounts to replace (5.7) by:

$$\inf_{\sigma\in\mathcal{M}(\Omega,\mathbb{R}^d)}\Big\{\int_\Omega L\Big(x,\frac{\mathrm{d}\sigma}{\mathrm{d}|\sigma|}(x)\Big)\mathrm{d}|\sigma|(x)\ :\ -\mathrm{div}(\sigma)=f\Big\}. \tag{5.9}$$

where $|\sigma|$ is the total variation measure of the vector-valued measure σ and $\frac{\mathrm{d}\sigma}{\mathrm{d}|\sigma|}$ is the density of σ with respect to $|\sigma|$. It is then obvious by (5.3) and Banach-Alaoglu Theorem that the relaxed problem (5.9) admits solutions. To sum up, we have the following duality and attainment relations

$$\mathrm{MK}(L,f):=\min(5.1)=\max(5.4)=\inf(5.7)=\min(5.9) \tag{5.10}$$

where we have denoted $\mathrm{MK}(L,f)$ the common value of (5.1), (5.4) and (5.7).

2.2 Relations between the three problems

We now discuss in a slightly formal way, relationships between the three problems. For further use, let us denote by $B(x)$ and $B^*(x)$ respectively the unit ball for $L(x,.)$ and $L^*(x,.)$:

$$B(x):=\{\sigma\in\mathbb{R}^d,\ L(x,\sigma)\le 1\},\ B^*(x):=\{q\in\mathbb{R}^d,\ :\ L^*(x,q)\le 1\}$$

and recall

$$L(x,\sigma)L^*(x,q)\ge\sigma\cdot q,\ L(x,\sigma)=\sup_{q\in B^*(x)}q\cdot\sigma,\ L^*(x,q)=\sup_{\sigma\in B(x)}q\cdot\sigma. \tag{5.11}$$

Recalling that if C is a closed convex subset of $\mathbb{R}^d$ and $z\in C$ the normal cone of C at z, $N_C(z)$ is by definition:

$$N_C(z):=\{\xi\in\mathbb{R}^d\ :\ \xi\cdot z\ge\xi\cdot y,\ \forall y\in C\}.$$

So if non zero vectors σ and q satisfy $L(x,\sigma)L^*(x,q)=q\cdot\sigma$ this exactly means

$$q\in N_{B(x)}\Big(\frac{\sigma}{L(x,\sigma)}\Big) \tag{5.12}$$

or equivalently

$$\sigma\in N_{B^*(x)}\Big(\frac{q}{L^*(x,q)}\Big). \tag{5.13}$$

In the case where $B(x)$ or $B^*(x)$ is smooth the normal cones at a point of $\partial B(x)$ or $\partial B^*(x)$ are simply the half line generated by the normal vectors (Gauss maps) and thus the previous relations give an unambiguous information on the relation between the direction of q and σ.

Any optimal plan π for (5.1) is related to any optimal potential u for (5.4) by the complementary slackness condition:

$$u(y) - u(x) = d_L(x, y), \ \forall (x, y) \in \operatorname{spt}(\pi). \tag{5.14}$$

Let then $(x, y) \in \operatorname{spt}(\pi)$ and let $t \in [0, 1] \mapsto \gamma_{x,y}(t)$ be a geodesic between x and y, it is easy to deduce from (5.14) and the fact that u is 1-d_L Lipschitz that one also has for every (s, t) such that $0 \leq s \leq t \leq 1$:

$$u(\gamma_{x,y}(t)) - u(\gamma_{x,y}(s)) = d_L(\gamma_{x,y}(t), \gamma_{x,y}(s)) = (t - s)d_L(x, y). \tag{5.15}$$

In other words, u somehow grows at maximal speed allowed by the Lipschitz constraint on the geodesic $\gamma_{x,y}$. If u was smooth we could further write:

$$u(y) - u(x) = \int_0^1 \nabla u(\gamma_{x,y}(s)) \cdot \dot{\gamma}_{x,y}(s) \mathrm{d}s$$

and then

$$\int_0^1 \nabla u(\gamma_{x,y}(s)) \cdot \dot{\gamma}_{x,y}(s) \mathrm{d}s = d_L(x, y) = \int_0^1 L(\gamma_{x,y}(s), \dot{\gamma}_{x,y}(s)) \mathrm{d}s.$$

but since $L^*(\gamma_{x,y}(s), \nabla u(\gamma_{x,y}(s))) \leq 1$ we pointwise have $\nabla u(\gamma_{x,y}(s)) \cdot \dot{\gamma}_{x,y}(s) \leq L(\gamma_{x,y}(s), \dot{\gamma}_{x,y}(s))$ so that

$$\nabla u(\gamma_{x,y}(s)) \cdot \dot{\gamma}_{x,y}(s) = L(\gamma_{x,y}(s), \dot{\gamma}_{x,y}(s)), \ \forall s \in [0, 1], \tag{5.16}$$

which also gives

$$L^*(\gamma_{x,y}(s), \nabla u(\gamma_{x,y}(s))) = 1, \ \forall s \in [0, 1]. \tag{5.17}$$

This expresses in a local way the fact that the Lipschitz constraint on u is binding on geodesics (this is again formal). Note that (5.17) gives a precise relation between $\nabla u(x)$ and the direction of geodesics passing through x: they are tangent to a vector in the normal cone $N_{B^*(x)}(\nabla u(x))$.

Now if σ solves (5.7) and u is a solution of (5.4), then complementary slackness takes the form

$$L(x, \sigma(x)) = \sigma(x) \cdot \nabla u(x) \text{ a.e.} \tag{5.18}$$

hence

$$\sigma(x) \neq 0 \Rightarrow L^*(x, \nabla u(x)) = 1, \tag{5.19}$$

which again expresses that the Lipschitz constraint is binding on the support of the transport density. The direction of optimal flows and gradients of Kantorovich potentials are therefore related by the duality relations

$$\sigma(x) \neq 0 \Rightarrow \sigma(x) \in N_{B^*(x)}(\nabla u(x)), \ \nabla u(x) \in N_{B(x)}\Big(\frac{\sigma(x)}{L(x, \sigma(x))}\Big). \tag{5.20}$$

It remains to investigate the relations between optimal plans and optimal flow fields. The following (heuristic) construction is well-known (see for instance [7]) in the euclidean setting: let π be an optimal plan i.e. a solution for (5.1). For every

$(x,y)\in \mathrm{spt}(\pi)$ let $t\in[0,1]\mapsto \gamma_{x,y}(t)$ be a geodesic between x and y and define the vector-valued measure σ_π by

$$\int_{\overline{\Omega}} F\mathrm{d}\sigma_\pi := \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 F(\gamma_{x,y}(s))\cdot\dot{\gamma}_{x,y}(s)\mathrm{d}s\Big)\mathrm{d}\pi(x,y) \tag{5.21}$$

for every $F\in C(\overline{\Omega},\mathbb{R}^d)$. Let $\phi\in C^1(\overline{\Omega})$, using the fact that $\pi\in\Pi(f^-,f^+)$ then gives

$$\int_{\overline{\Omega}} \nabla\phi\mathrm{d}\sigma_\pi = \int_{\overline{\Omega}\times\overline{\Omega}} (\phi(y)-\phi(x))\mathrm{d}\pi(x,y) = \langle\phi,f\rangle$$

i.e. $-\mathrm{div}(\sigma_\pi)=f$ so that σ_π is admissible for the minimal flow problem (5.7). To see that σ_π is actually optimal, we consider a Kantorovich potential i.e. a solution u of (5.4). Thanks to (5.10), it is enough to show that

$$\int_\Omega L(x,\sigma_\pi(x))\mathrm{d}x \le \langle u,f\rangle.$$

On the one hand, observing that:

$$\int_\Omega L(x,\sigma_\pi(x))\mathrm{d}x = \sup\Big\{\int_\Omega F(x)\cdot\sigma_\pi(x)\mathrm{d}x \ : \ L^*(x,F(x))\le 1\Big\}$$

and that if $L^*(x,F(x))\le 1$ then

$$\begin{aligned}\int_\Omega F(x)\cdot\sigma_\pi(x)\mathrm{d}x &= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 F(\gamma_{x,y}(s))\cdot\dot{\gamma}_{x,y}(s)\mathrm{d}s\Big)\mathrm{d}\pi(x,y)\\ &\le \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 L(\gamma_{x,y}(s),\dot{\gamma}_{x,y}(s))\mathrm{d}s\Big)\mathrm{d}\pi(x,y)\end{aligned}$$

we get

$$\int_\Omega L(x,\sigma_\pi(x))\mathrm{d}x \le \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 L(\gamma_{x,y}(s),\dot{\gamma}_{x,y}(s))\mathrm{d}s\Big)\mathrm{d}\pi(x,y).$$

On the other hand, thanks to the complementary slackness condition (5.16) and $-\mathrm{div}(\sigma_\pi)=f$, we have

$$\begin{aligned}\langle u,f\rangle = \int_\Omega \nabla u\cdot\sigma_\pi &= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 \nabla u(\gamma_{x,y}(s))\cdot\dot{\gamma}_{x,y}(s)\mathrm{d}s\Big)\mathrm{d}\pi(x,y)\\ &= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\int_0^1 L(\gamma_{x,y}(s)),\dot{\gamma}_{x,y}(s))\mathrm{d}s\Big)\mathrm{d}\pi(x,y).\end{aligned}$$

This proves the optimality of σ_π.

2.3 Lagrangian and saddle-point

Rewrite (5.4) as

$$\inf_{u,q}\Big\{-\langle u,f\rangle + G(q) \ : \ q=\nabla u \text{ a.e.}\Big\}$$

where

$$G(q) := \int_\Omega \mathcal{G}(x,q(x))\mathrm{d}x$$

and

$$\mathcal{G}(x,q) := \begin{cases} 0, & \text{if } L^*(x,q) \le 1 \\ +\infty, \text{otherwise} \end{cases}$$

and then rewrite (5.4)-(5.7) as the saddle point problem

$$\inf_{u,q} \sup_{\sigma} \mathcal{L}(u,q,\sigma)$$

where the Lagrangian $\mathcal{L}$ is defined by

$$\begin{aligned} \mathcal{L}(u,q,\sigma) := -\langle u, f\rangle + \int_\Omega \mathcal{G}(x,q(x))\mathrm{d}x \\ + \int_\Omega \sigma(x) \cdot (\nabla u(x) - q(x))\mathrm{d}x. \end{aligned}$$

For $r > 0$, let us also introduce the augmented Lagrangian

$$\begin{aligned} \mathcal{L}_r(u,q,\sigma) := -\langle u, f\rangle + \int_\Omega \mathcal{G}(x,q(x))\mathrm{d}x \\ + \int_\Omega \sigma(x) \cdot (\nabla u(x) - q(x))\mathrm{d}x + \frac{r}{2}\int_\Omega |\nabla u - q|^2. \end{aligned}$$

Recall that $\mathcal{L}$ and $\mathcal{L}_r$ have the same saddle-points (see [59, 61]). Note that in both $\mathcal{L}$ and $\mathcal{L}_r$ we multiply the L^∞ vector field ∇u by σ, which a priori only makes sense only if σ is L^1. Existence of saddle-points is therefore not guaranteed unless there is an L^1 solution to (5.7). However at the level of the discretized problems (see next section), there is no such regularity issue, there exists saddle-points for the discretized Lagrangian and finding such saddle-points is equivalent to solving (5.4) and (5.7) simultaneously.

3 Discretization and algorithm

3.1 Discretization

We now consider suitable approximations of our problems by finite-dimensional (convex) ones using finite elements. In these finite dimensional-approximations existence of saddle-points is not an issue anymore. More precisely, consider a family of regular triangulations $\mathcal{T}_h$ of the domain (which we now assume to be two-dimensional) indexed by the typical meshsize h (i.e. the diameter of each $T \in \mathcal{T}_h$ is less than Ch for some positive constant C), let $E_h \subset W^{1,\infty}(\Omega)$ be the corresponding finite-dimensional space of Lagrange P_1 (piecewise linear) finite elements of order 1 (a similar analysis can be done for higher order finite-elements) whose generic elements are denoted u_h. Slightly abusing notations, we shall consider u_h both as a finite-dimensional vector and a Lipschitz, piecewise linear function defined on the whole domain, the gradient of u_h has piecewise constant components, it is still denoted ∇u_h. We further assume that the mesh is regular in the sense that the Lagrange interpolate map $I_h : W^{1,\infty}(\Omega) \to E_h$ satisfies

$$\lim_{h\to 0} \|\nabla v - \nabla(I_h(v))\|_{L^\infty} \to 0, \ \forall v \in C^1(\overline{\Omega}). \tag{5.22}$$

We also approximate the linear form f by $f_h \in (E_h)^* \simeq E_h$ (again with $\langle f_h, 1\rangle = 0$) in such a way that f_h weakly converges to f in the sense of measures as $h \to 0$.

We then consider the approximation of (5.4):

$$\sup_{u_h \in K_h} \langle f_h, u_h \rangle \tag{5.23}$$

where K_h is the convex subset of E_h consisting of all u_h's in E_h such that for every $T \in \mathcal{T}_h$ one has

$$L^*(x_T, \nabla u_h|_T) \leq 1, \tag{5.24}$$

where x_T is a given point in T (for instance its center of mass or one of its vertices). To prove that this is a consistent approximation of Kantorovich problem (5.4), it is useful to observe first that smooth functions are dense in the admissible set for (5.4):

Lemma 5.1. *Let u be a 1-d_L function, then there exists a sequence u_n of $C^\infty(\mathbb{R}^d)$, 1-d_L-functions converging uniformly on $\overline{\Omega}$ to u. In particular this implies that*

$$\max(5.4) = \sup\Big\{\langle u, f\rangle : \ u \textit{ is } 1\textit{-Lipschitz for } d_L \textit{ and } C^\infty(\mathbb{R}^d)\Big\}. \tag{5.25}$$

Proof. First extend u on the whole of $\mathbb{R}^d$ by setting

$$u(x) = \inf_{y \in \overline{\Omega}} \{u(y) + d_L(x, y)\}.$$

Consider a standard mollifying kernel $\rho_\varepsilon(x) := \varepsilon^{-d}\rho(\varepsilon^{-1}x)$ with $\rho \in C_c^\infty(\mathbb{R}^d)$, $\rho \geq 0$, $\rho(x) = 0$ for $|x| \geq 1$ and $\int_{\mathbb{R}^d} \rho = 1$. Then for $\varepsilon > 0$ and $\delta > 0$, define then the smooth function $u_{\varepsilon,\delta} := \frac{1}{1+\delta}\rho_\varepsilon \star u$. We then have for every $x \in \overline{\Omega}$, using the convexity of $L^*(x, .)$, Jensen's inequality and the fact that $L^*(y, \nabla u(y)) \leq 1$ a.e.

$$\begin{aligned}
L^*(x, \nabla u_{\varepsilon,\delta}(x)) &\leq \frac{1}{1+\delta}\int_{\mathbb{R}^d} \rho_\varepsilon(x-y) L^*(x, \nabla u(y))\mathrm{d}y \\
&= \frac{1}{1+\delta}\int_{\mathbb{R}^d} \rho_\varepsilon(x-y) L^*(y, \nabla u(y))\mathrm{d}y \\
&+ \frac{1}{1+\delta}\int_{\mathbb{R}^d} \rho_\varepsilon(x-y)(L^*(x, \nabla u(y)) - L^*(y, \nabla u(y)))\mathrm{d}y \\
&\leq \frac{1+\omega(\varepsilon)}{1+\delta}
\end{aligned}$$

where

$$\omega(\varepsilon) := \sup |\{L^*(x,q) - L^*(y,q)|, \ x \in \overline{\Omega}, \ |x-y| \leq \varepsilon, \ |q| \leq \|\nabla u\|_{L^\infty}\}.$$

Thus $u_{\varepsilon,\delta}$ is 1-d_L Lipschitz as soon as $\omega(\varepsilon) \leq \delta$, this clearly proves the desired result since L^* is continuous.

□

One easily deduces the following convergence result:

Proposition 5.1. *Let u_h be a solution of (5.23) normalized so as to have zero mean, then for some vanishing sequence of meshsizes $h_n \to 0$ as $n \to \infty$, u_{h_n} converges uniformly to some Kantorovich potential u i.e. some solution of (5.4).*

Proof. Thanks to (5.3), u_h is uniformly Lipschitz, since it has nonzero mean, thanks to Ascoli's theorem, for some vanishing sequence of meshsizes, it converges in $C(\overline{\Omega})$ to some Lipschitz function u. Thanks to Banach-Alaoglu's Theorem, we may also assume that ∇u_h converges weakly $*$ in L^∞ to ∇u. To check that u is 1-d_L Lipschitz, it is enough to show that for every $\sigma \in C(\overline{\Omega}, \mathbb{R}^2)$ one has

$$\int_\Omega \sigma \cdot \nabla u \leq \int_\Omega L(x, \sigma(x))\mathrm{d}x. \tag{5.26}$$

Since $u_h \in K_h$ we have for every $T \in \mathcal{T}_h$, $\nabla u_h|_T \cdot \sigma(x_T) \leq L(x_T, \sigma(x_T))$, multiplying by the measure of T, summing over all triangles of $\mathcal{T}_h$ and letting $h \to 0$ gives (5.26). It remains to prove that u solves (5.4), which thanks to Lemma Lemma 5.1 amounts to show that $\langle f, u\rangle \geq \langle f, v\rangle$ for every smooth and 1-d_L-Lipschitz function v. Let then v be such a smooth and 1-d_L-Lipschitz function, for every $T \in \mathcal{T}_h$, we have

$$\begin{aligned} L^*(x_T, \nabla I_h(v)(x_T)) &= L^*(x_T, \nabla v(x_T)) + L^*(x_T, \nabla I_h(v)(x_T)) - L^*(x_T, \nabla v(x_T)) \\ &\leq 1 + \omega_h \end{aligned}$$

where

$$\omega_h := \sup_{T \in \mathcal{T}_h} |L^*(x_T, \nabla I_h(v)(x_T)) - L^*(x_T, \nabla v(x_T))| \leq C\|\nabla v - \nabla(I_h(v))\|_{L^\infty}$$

tends to 0 as $h \to 0$ thanks to (5.22). Then defining $v_h := (1+\omega_h)^{-1} I_h(v)$, we have $v_h \in K_h$ and v_h converges uniformly to v as $h \to 0$, passing to the limit in $\langle f_h, u_h\rangle \geq \langle f_h, v_h\rangle$, we can conclude that u is a Kantorovich potential.

□

3.2 Augmented Lagrangian algorithm

From now on, we drop the dependence in h in the approximation parameter and slightly abusing notation, we keep the same notations as in the continuous framework, eventhough in what follows we actually consider the discretization of the augmented Lagrangian $\mathcal{L}_r$. Existence of a saddle-point is not an issue at the level of the finite-dimensional approximation and convergence of the augmented Lagrangian algorithm recalled below is well-known (see Eckstein and Bertsekas [53]).

The augmented Lagrangian algorithm ALG2 splitting scheme, consists, starting from $(u^0, q^0, \sigma^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ to generate inductively a sequence (u^k, q^k, σ^k) as follows (abusing notations we still denote by ∇ the discretization of the gradient):

— **Step 1:** minimization with respect to u:

$$u^{k+1} := \mathrm{argmin}_{u \in \mathbb{R}^n} \Big\{ -\langle u, f\rangle + \sigma^k \cdot \nabla u + \frac{r}{2}|\nabla u - q^k|^2 \Big\}, \tag{5.27}$$

— **Step 2:** minimization with respect to q:

$$q^{k+1} := \mathrm{argmin}_{q \in \mathbb{R}^m} \Big\{ G(q) - \sigma^k \cdot q + \frac{r}{2}|\nabla u^{k+1} - q|^2 \Big\}, \tag{5.28}$$

— **Step 3:** update the multiplier by the gradient ascent formula

$$\sigma^{k+1} = \sigma^k + r(\nabla u^{k+1} - q^{k+1}). \tag{5.29}$$

Step 1 consists in solving a Laplace equation :

$$-r(\Delta u^{k+1} - \operatorname{div}(q^k)) = f + \operatorname{div}(\sigma^k) \text{ in } \Omega, \tag{5.30}$$

together with the Neumann boundary condition

$$r\frac{\partial u^{k+1}}{\partial \nu} = rq^k \cdot \nu - \sigma^k \cdot \nu \text{ on } \partial\Omega. \tag{5.31}$$

Step 2 is a pointwise projection problem

$$q^{k+1}(x) = p_{B^*(x)}\Big(\nabla u^{k+1} + \frac{\sigma^k}{r}\Big),$$

where $p_{B^*(x)}$ is the projection onto $B^*(x) := \{q \in \mathbb{R}^d \ : \ L^*(x,q) \leq 1\}$ the unit ball for $L^*(x,.)$.

3.3 Examples

We now give some details on how to perform the projection step 2 in practice. For the sake of simplicity we shall here drop the dependence of L and L^* in x.

The Riemannian case

In the Riemannian case $L(v) = (Av \cdot v)^{\frac{1}{2}}$ for some symmetric positive definite matrix A. Up to diagonalizing A, there is no loss of generality in assuming that $L(v) = (\sum_{i=1}^d \lambda_i v_i^2)^{\frac{1}{2}}$ with $\lambda_i > 0$ the eigenvalues of A. The dual norm L^* is then given by $L^*(q) = (\sum_{i=1}^d \lambda_i^{-1} q_i^2)^{\frac{1}{2}}$. The projection p_{B^*} onto $B^* := \{q \in \mathbb{R}^d \ : \ L^*(q) \leq 1\}$ is almost explicit:

$$p_B^*(\overline{q}) = \begin{cases} \overline{q}, \text{ if } \overline{q} \in B^*, \\ \left(\frac{\lambda_1 \overline{q}_1}{\lambda_1+\alpha}, \cdots \frac{\lambda_d \overline{q}_d}{\lambda_d+\alpha}\right) \text{ with } \alpha \text{ the unique positive root of (5.32) otherwise.} \end{cases}$$

where the nonlinear equation to be solved by α reads

$$1 = \sum_{i=1}^d \frac{\lambda_i \overline{q}_i^2}{(\lambda_i + \alpha)^2}. \tag{5.32}$$

This single equation is monotone in α and can be efficiently solved by Newton's method.

The case where $L(x,.)$ is defined by finitely directions

The second case we have in mind is the polyhedral case where L is defined by finitely many directions. More precisely (and again this is for a fixed x), we are given a collection of unit vectors $v_1, \ldots, v_k$ which we complete by $v_{k+1} = -v_1, \ldots, v_{2k} =$

$-v_k$ and such that 0 belongs to the interior of the (symmetric) convex polytope $\mathrm{co}(\{v_j,\ j=1,\dots,2k\})$. We are also given positive reals $(\xi_j)_{j=1,\dots,2k}$ with $\xi_{j+k}=\xi_j$ for $j\in\{1,\dots,k\}$ and then consider the crystalline norm

$$L(v) := \inf\Big\{\sum_{j=1}^{2k}\xi_j\alpha_j\ :\ \sum_{j=1}^{2k}\alpha_j v_j = v\Big\}.$$

It is immediate to see that L is the gauge of the symmetric convex polytope $B := \mathrm{co}(\{\xi_j^{-1}v_j,\ j=1,\dots,2k\})$ which is then its unit ball. The dual norm L^* is then explicitly given by

$$L^*(q) := \max_{j=1,\dots,k}\xi_j^{-1}|q\cdot v_j|, \tag{5.33}$$

its dual unit ball B^* is then defined by the inequalities $|q\cdot v_j|\le\xi_j$ for $j=1,\dots,k$. In dimension two, the projection onto B^* can be easily performed as follows. First, compute the vertices and sides of B^* (note that the latter have one of the vectors v_j as normal, so these computations can be done in an automatic way) so as to be able to represent $B^*=\mathrm{co}(\{S_i,\quad i=1,\dots,2l\})$ where $S_1,\dots,S_l$ are the successive vertices of B^* and denote by ν_i the unit exterior normal to the side $[S_i,S_{i+1}]$. Now if $\overline{q}$ is a generic vector of the plane belonging to the complement of B^* (otherwise its projection is $\overline{q}$), then $\overline{q}$ belongs either to one half strip $[S_i,S_{i+1}]+\mathbb{R}_+\nu_i$ and in this case its projection on B^* coincides with its projection on the line $S_i+\nu_i^\perp$ or it belongs to one of the sectors $S_i+\mathbb{R}_+\nu_{i-1}+\mathbb{R}_+\nu_i$ and in this case the projection of $\overline{q}$ is the vertex S_i.

We illustrate with the following example : $k=4$,

$$v_j=\left(\cos\left(\frac{(j-1)\pi}{k}\right),\sin\left(\frac{(j-1)\pi}{k}\right)\right),\xi_1=2.5,\xi_2=2,\xi_3=1.5\text{ and }\xi_4=3.$$

In fact the vector v_4 is useless since ξ_4 is very large with respect to the other ones. So the ball B^* is only defined by the inequalities $|q\cdot v_j|\le\xi_j$ for $j=1,\dots,3$.

The point $\overline{q}_1$ is in the half strip $[S_6,S_1]+\mathbb{R}_+v_1$ so that its projection is on the segment $[S_6,S_1]$. The point $\overline{q}_2$ belongs to the sector $S_6+\mathbb{R}_+v_6+\mathbb{R}_+v_1$ so that its projection is S_6.

4 Results

We use the software FreeFem++ (see [68]) to implement the numerical ALG2 scheme described above. The Lagrangian finite elements and notations used in Subsection 3.2 are taken here. We use P_2 FE for u_h and P_1 for (q_h,σ_h) (approximation is better and convergence is faster than with P_0 and P_1). As emphasized in the previous subsection the first step and the third one are always the same. Only the projection step 2 changes according to the geometry of the Finsler metric. For our numerical simulations Ω is a $2D$ square ($x=(x_1,x_2)\in[0,1]^2$) and we test with different f :

$$f_1^- := e^{-40*((x_1-0.75)^2+(x_2-0.25)^2}\text{ and }f_1^+ := e^{-40*((x_1-0.25)^2+(x_2-0.65)^2)},$$
$$f_2^- := e^{-40*((x_1-0.5)^2+(x_2-0.15)^2)}\text{ and }f_2^+ := e^{-40*((x_1-0.5)^2+(x_2-0.75)^2)}.$$

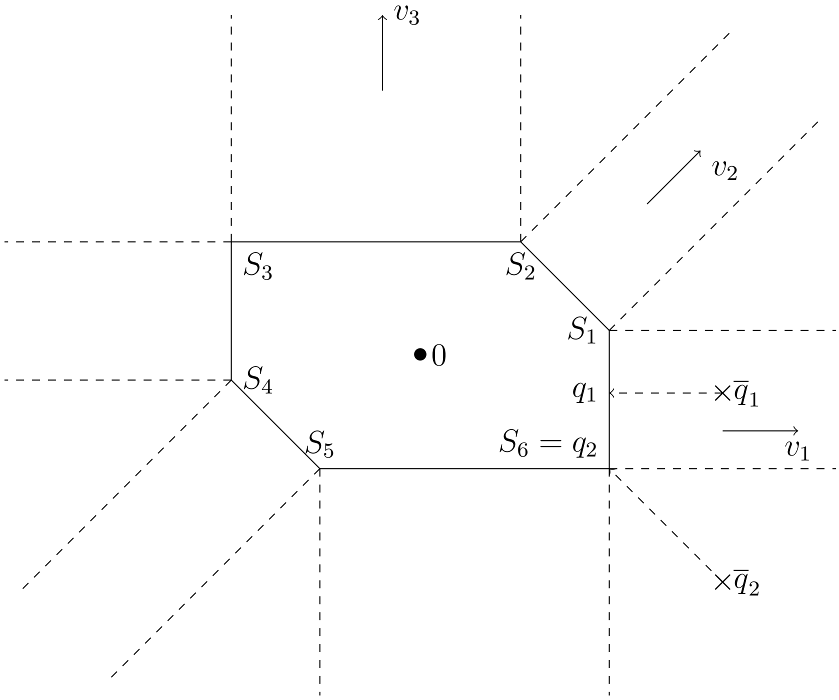

Figure 5.1 – An illustration of projection on a polyhedron.

In the third case, we take f_3^- a constant density and f_3^+ is the sum of three concentrated Gaussians

$$f_3^+(x_1,x_2) = e^{-400*((x-0.25)^2+(y-0.75)^2)} + e^{-400*((x-0.35)^2+(y-0.15)^2)} + e^{-400*((x-0.85)^2+(y-0.7)^2)}.$$

In the following two subsections, in each figure, there are two images. The top one represents σ and the bottom one corresponds to the level lines of u.

4.1 Riemannian case

Here $L(x,v) = (\lambda_1(x)v_1^2 + \lambda_2(x)v_2^2)^{\frac{1}{2}}$, we take one λ_i constant and the other one non constant, that is equal to the inverse of

$$g(x_1,x_2) = 1.5 - \exp(-100 * ((x_1-0.5)^2 + (x_2-0.5)^2)).$$

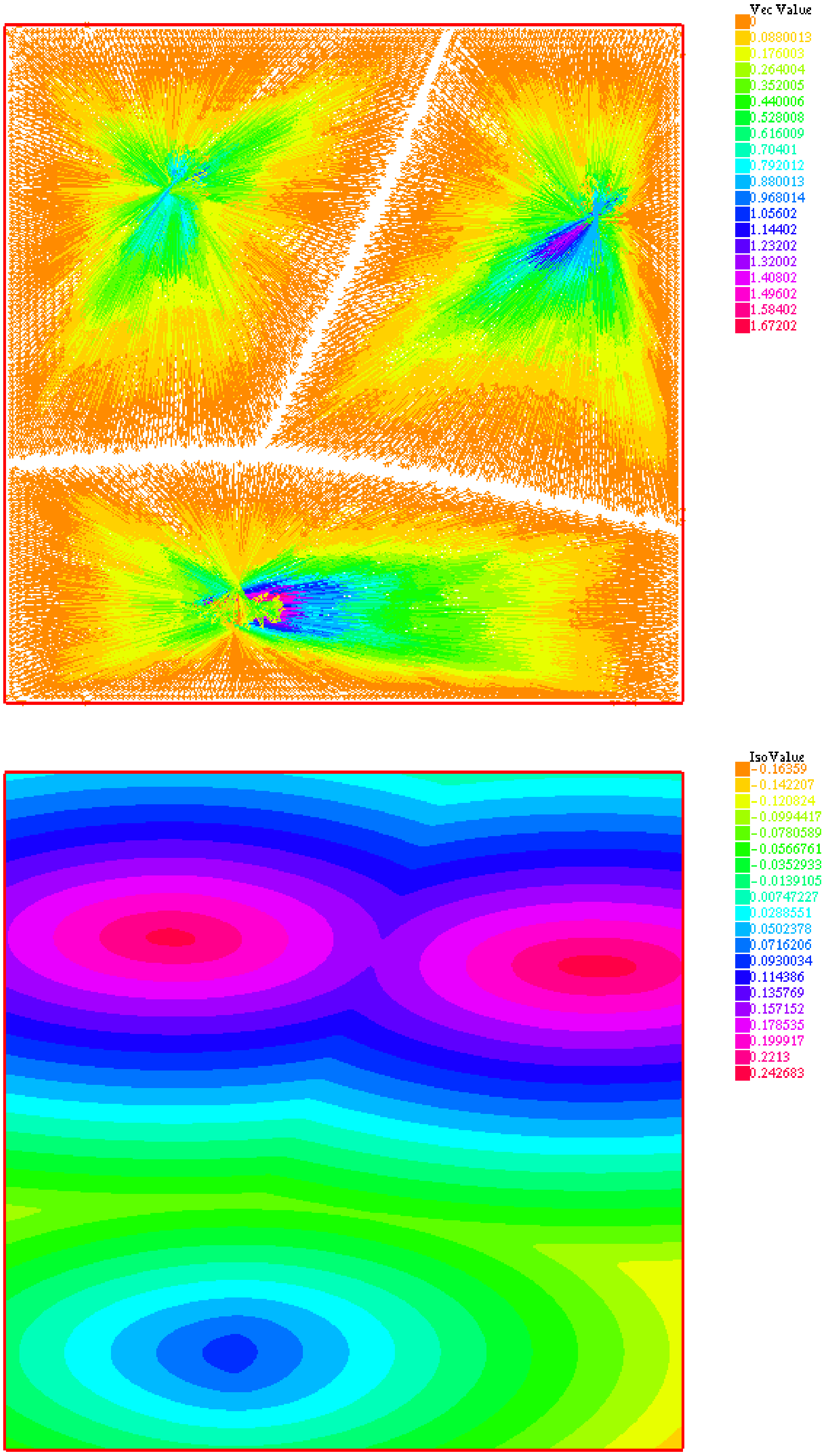

Figure 5.2 – Test case 1 : $\lambda_1 = 0.1$ and $\lambda_2 = 1/g$ with $f = f_3$.

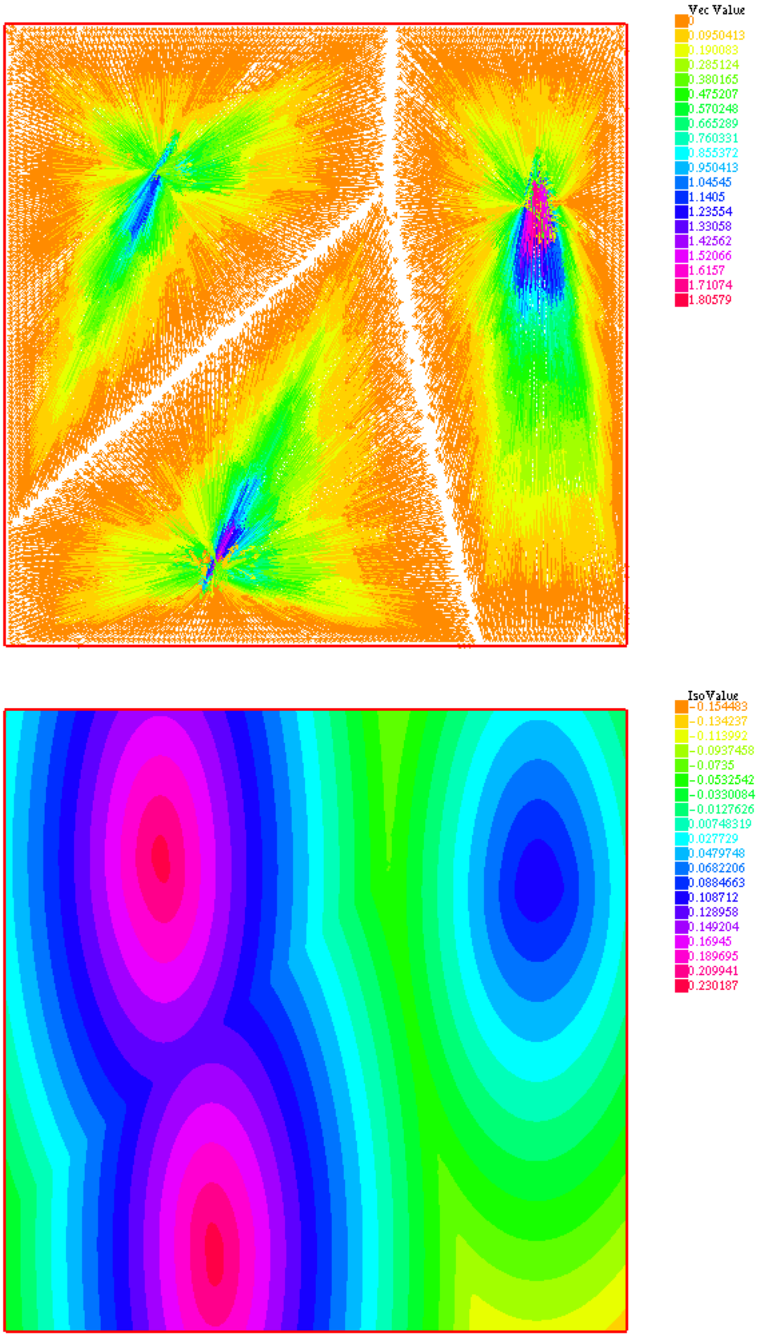

Figure 5.3 – Test case 2 : $\lambda_1 = 1/g$ and $\lambda_2 = 0.1$ with $f = f_3$.

4.2 Polyhedral case

We tested the following polyhedral examples. In Figure 5.6 we take $k = 2$ and v_1 perpendicular to v_2, the dual unit ball B^* then is a rectangle. In all other examples, $k = 15$ and the angle between two consecutive directions is π/k. The form of B^* then depends on the chosen ξ_j's. If the ξ_j's are (almost) equal, B^* is a polyhedron with thirty edges. It is in particular the case for Figure 5.5, Figure 5.8, Figure 5.9

and Figure 5.10. In the last examples, we have

$$\xi_j = \cos\left(\frac{2(j-1)\pi}{k}\right).$$

The ball B^* then has only 12 edges.

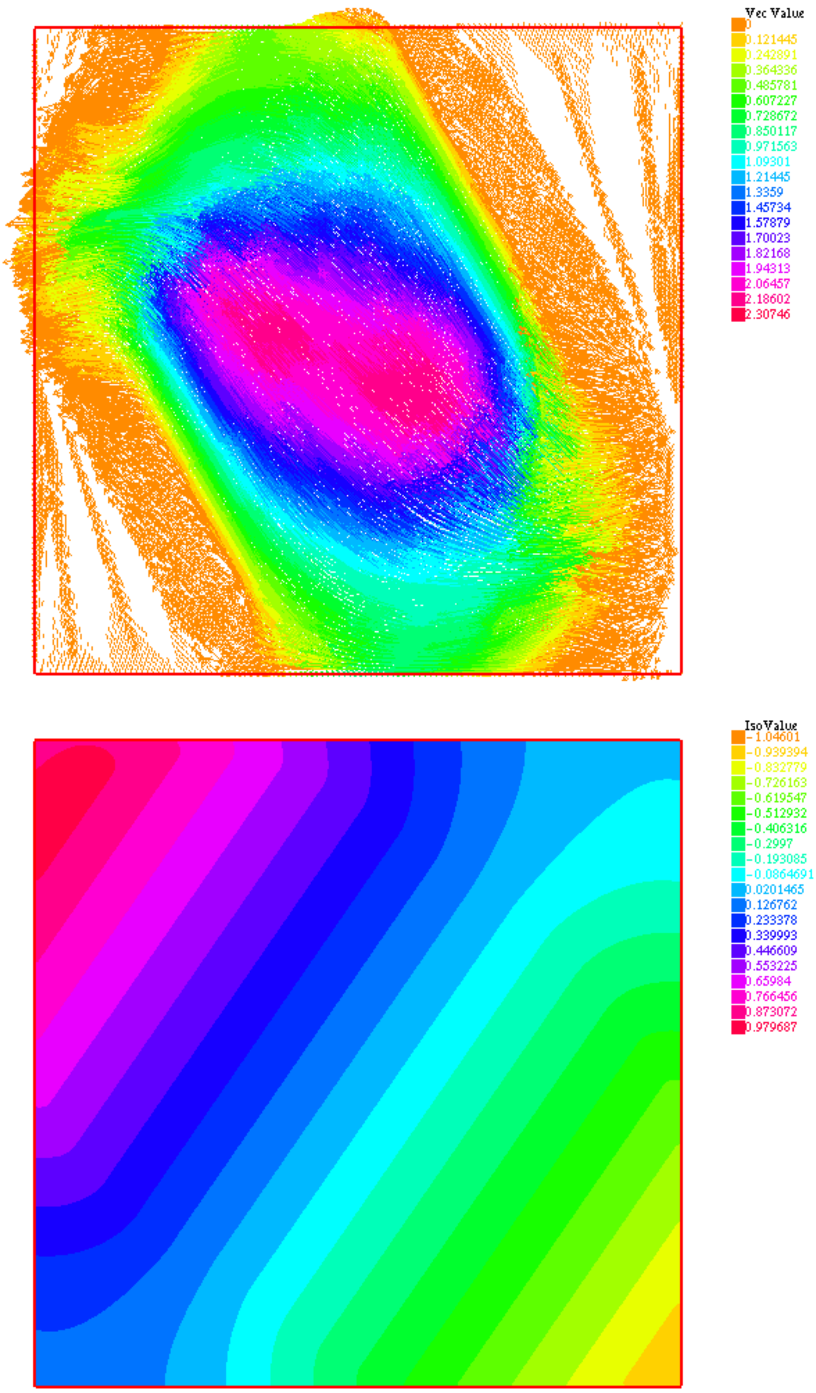

Figure 5.4 – Test case 3 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = \cos\left(\frac{2(j-1)\pi}{k}\right)$ for $j = 1, \ldots, k$ with $f = f_1$.

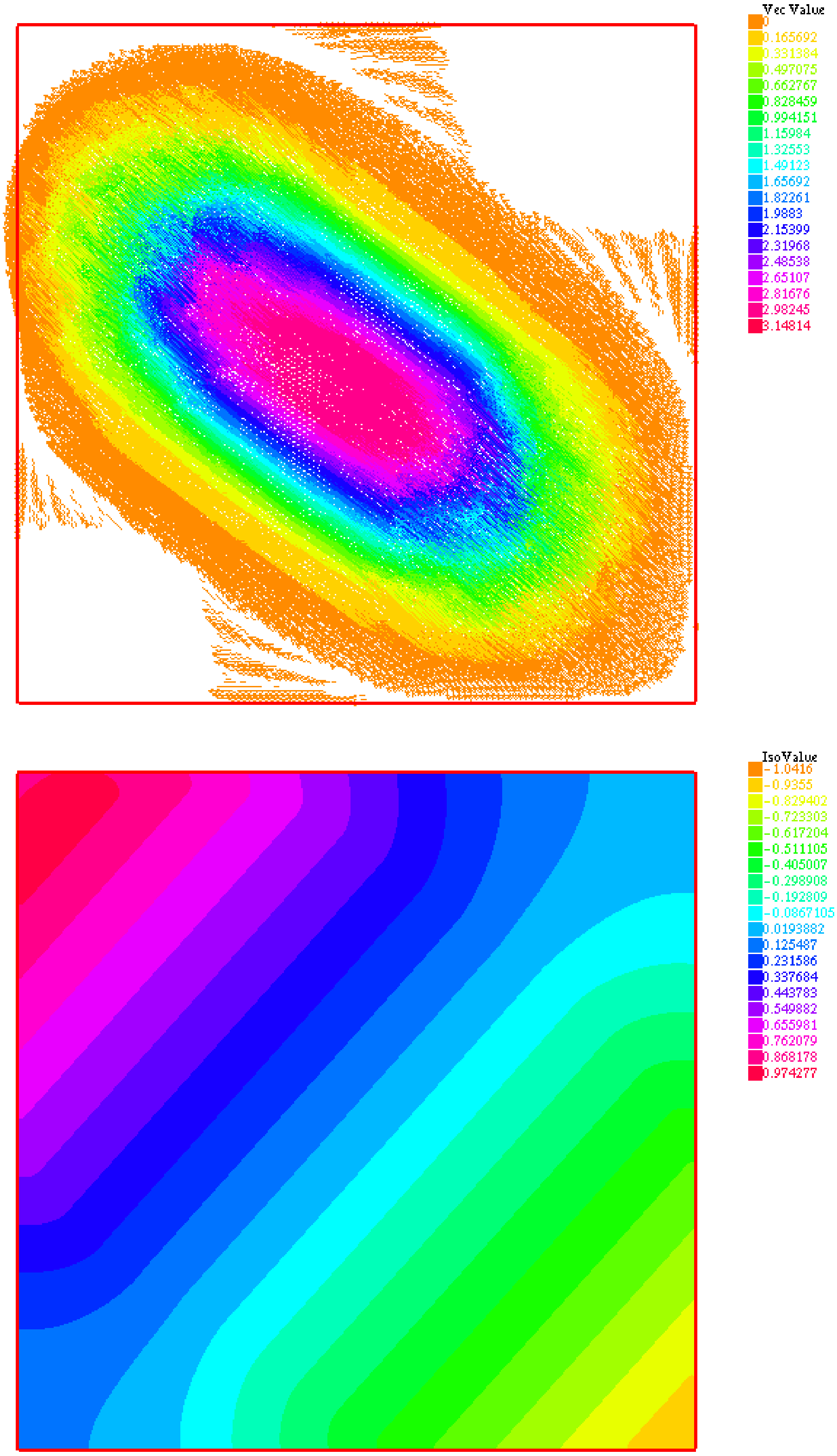

Figure 5.5 – Test case 4 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = 1.5$ for $j = 1, \ldots, k$ with $f = f_1$.

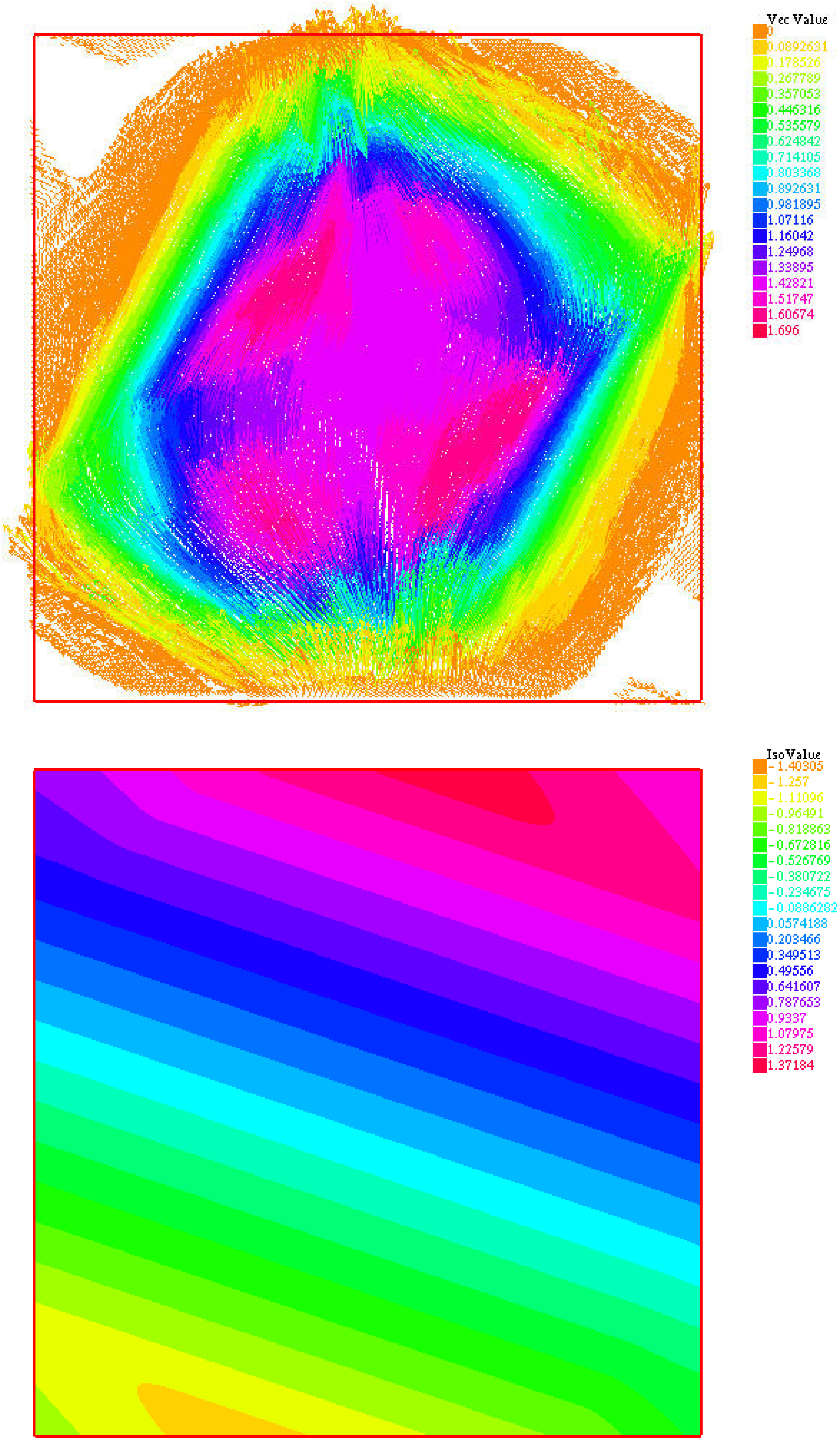

Figure 5.6 – Test case 5 : $k = 2$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k} + \frac{\pi}{3}\right), \sin\left(\frac{(j-1)\pi}{k} + \frac{\pi}{3}\right)\right)$ and $\xi_j = \cos\left(\frac{2(j-1)\pi}{k}\right)$ for $j = 1, \ldots, k$ with $f = f_2$.

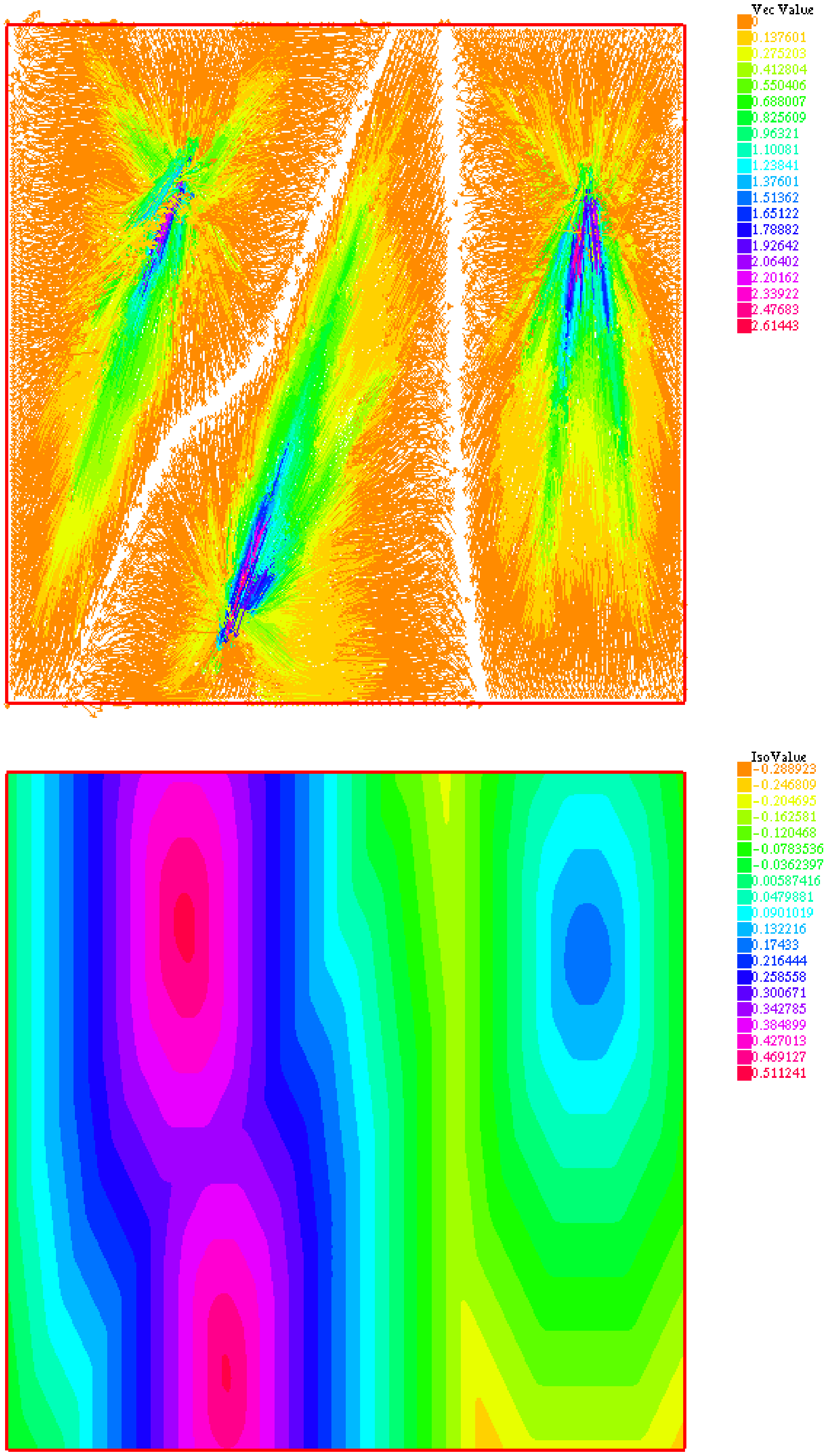

Figure 5.7 – Test case 6 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = \cos\left(\frac{2(j-1)\pi}{k}\right)$ for $j = 1, \ldots, k$ with $f = f_3$.

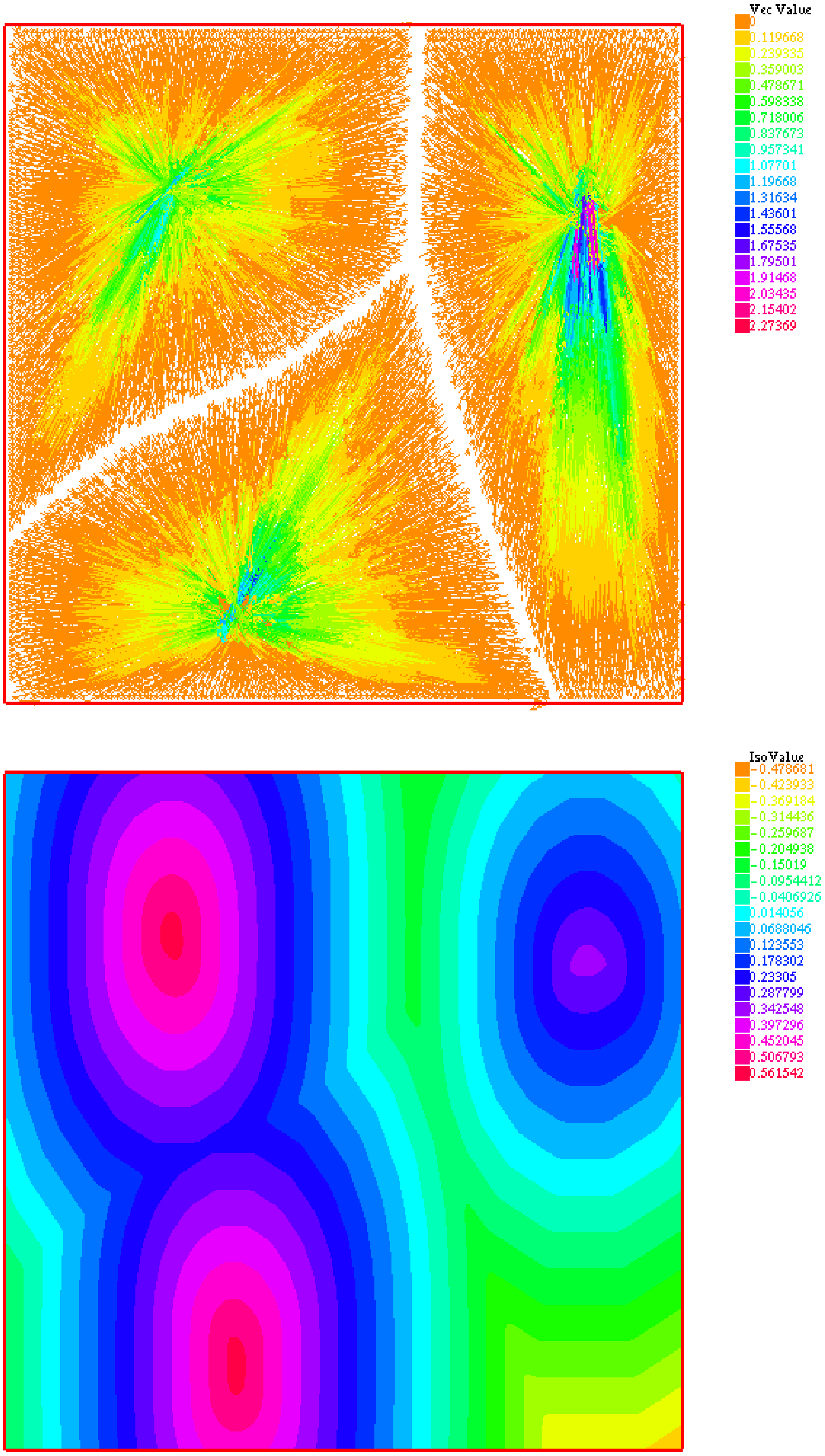

Figure 5.8 – Test case 7 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = \frac{1}{2}\cos\left(\frac{2(j-1)\pi}{k}\right)$ for $j = 1, \ldots, k$ with $f = f_3$.

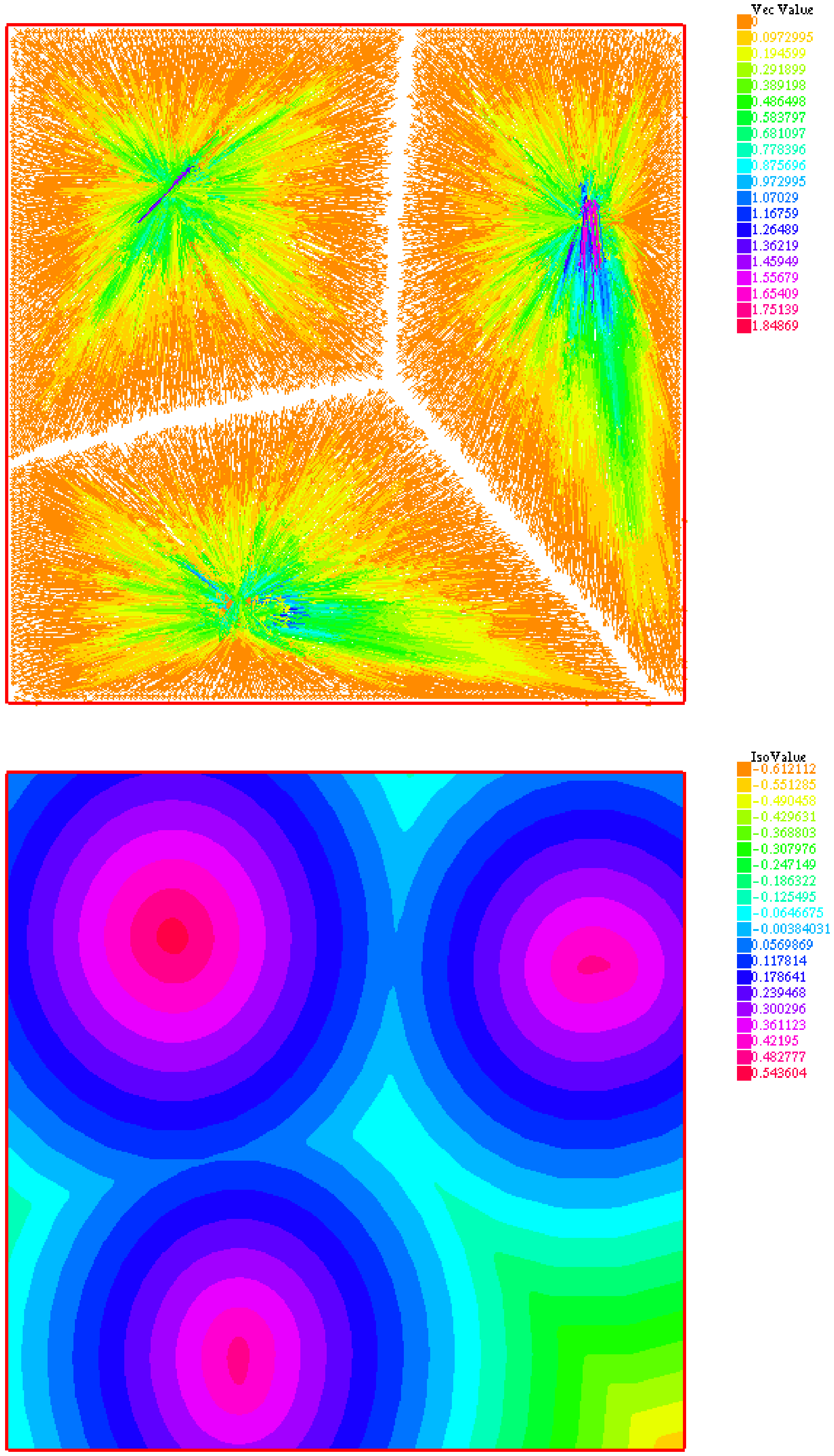

Figure 5.9 – Test case 8 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = \frac{1}{10}\cos\left(\frac{2(j-1)\pi}{k}\right)$ for $j = 1, \ldots, k$ with $f = f_3$.

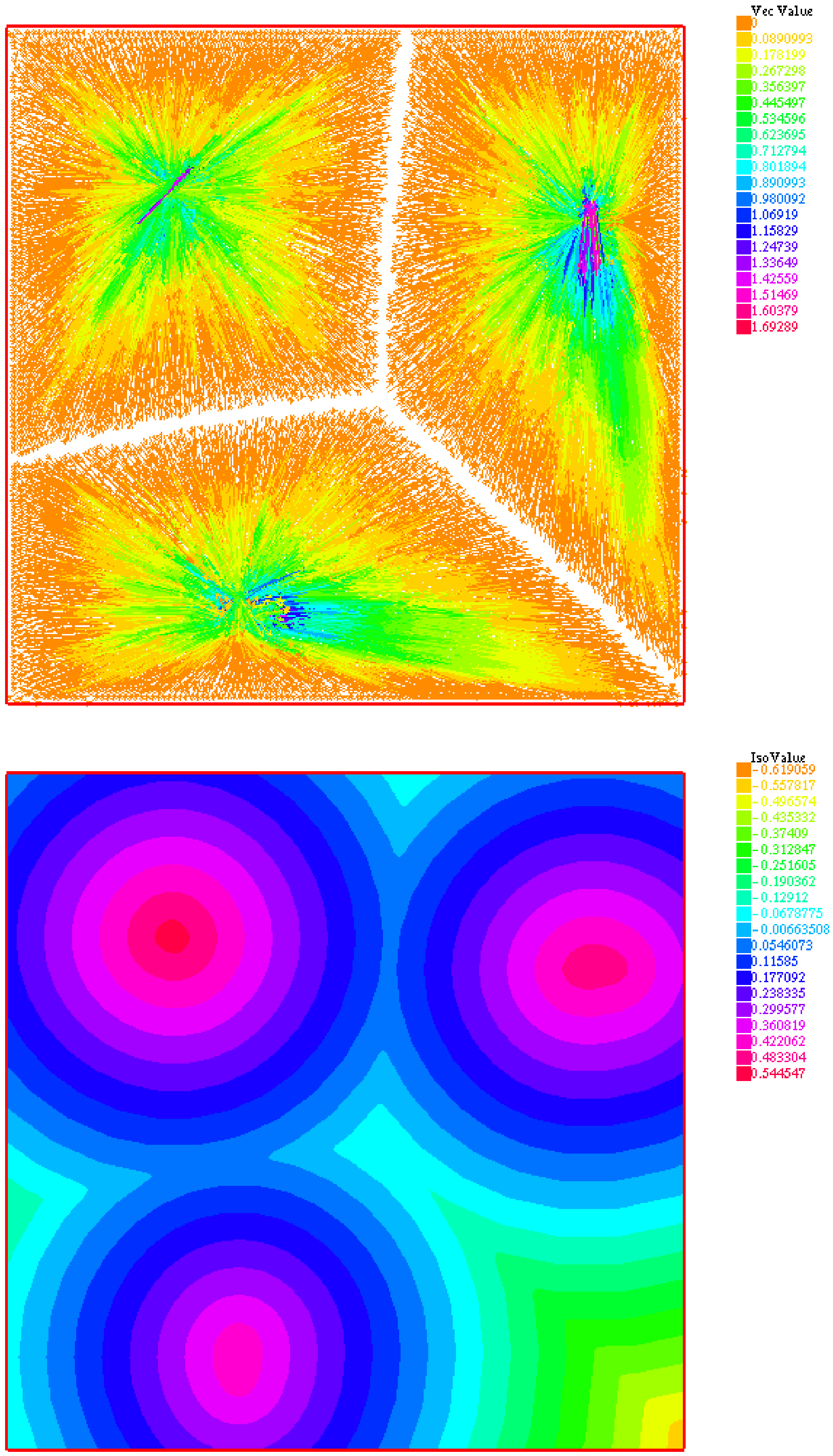

Figure 5.10 – Test case 9 : $k = 15$, $v_j = \left(\cos\left(\frac{(j-1)\pi}{k}\right), \sin\left(\frac{(j-1)\pi}{k}\right)\right)$ and $\xi_j = 1$ for $j = 1, \ldots, k$ with $f = f_3$.

4.3 Error Criteria

To analyze convergence in our simulations, we have considered three criteria corresponding to the optimality conditions:

$$- \operatorname{div}(\sigma) = f, \text{ in } \Omega \quad \sigma \cdot \nu = 0 \text{ on } \partial\Omega, \tag{5.34}$$

as well as duality relation

$$L(x, \sigma) = \sigma \cdot \nabla u \tag{5.35}$$

which can be equivalently rewritten in a dual way as

$$\sigma(x) \neq 0 \Rightarrow L^*(x, \nabla u(x)) = 1. \tag{5.36}$$

We use a triangulation of the unit square with $n = 1/h$ element on each side. We use the following convergence criteria:

1. DIV.Error $= \left(\int_{\Omega_h} (\operatorname{div}\sigma_h^k + f)^2\right)^{1/2}$ is the L^2 error on the divergence constraint.
2. BND.Error $= \left(\int_{\partial\Omega_h} (\sigma_h^k \cdot \nu)^2\right)^{1/2}$ is the $L^2(\partial\Omega_h)$ error on the Neumann boundary condition.
3. DUAL.Error $= \left(\int_{\Omega_h} |L(\cdot, \sigma_h^k(\cdot)) - \nabla u_h^k \cdot \sigma_h^k|\right)$ for the Riemannian case.
   DUAL.Error $= \left(\int_{\Omega_h} |L^*(\cdot, \nabla u_h^k(\cdot)) - 1|\chi_{\{|\sigma_h^k| > \varepsilon\}}\right)$ for the polyhedral case with $\varepsilon = 10^{-2}$.

The first two criteria represent the optimality conditions for the minimization of the Lagrangian with respect to u and the third one is for maximization with respect to σ. We do not take exactly the same criteria for both examples. Indeed, in the Riemannian case, $L(\cdot, \sigma)$ is simple to compute whereas in the polyhedral case, it is tedious in general. On the other hand, $L^*(\cdot, \nabla u(\cdot))$ has an explicit form given by (5.33).

We show the results of numerical simulations after 400 iterations for both cases.

Test case	DIV.Error	BND.Error	DUAL.Error	Time execution (seconds)
1	3.0940e-05	4.9502e-04	1.6274e-06	287
2	3.2576e-05	4.0942e-04	2.1978e-06	285
3	9.3806e-05	7.9803e-04	8.1512e-04	435
4	6.1646e-06	2.5572e-04	2.7813e-03	658
5	1.9829e-05	2.2784e-03	4.8522e-04	310
6	1.1407e-04	8.5331e-04	1.8588e-03	446
7	1.0402e-04	8.5816e-04	1.2846e-03	660
8	9.9358e-05	4.9236e-04	1.2181e-03	654
9	8.3469e-05	5.0099e-04	1.1265e-03	656

Table 5.1 – Convergence of the finite element discretization for all test cases.

Chapitre 6

Quelques perspectives

Nous donnons ici, de façon non exhaustive, quelques perspectives de problèmes assez directement liés aux travaux présentés dans cette thèse.

1 Modèles encore plus généraux

Dans le premier chapitre, nous avons fait des hypothèses structurelles pour obtenir le théorème 2.2. Nous sommes partis des cas cartésien puis hexagonal dans $\mathbb{R}^2$ pour construire ces modèles généraux. Une question naturelle est de déterminer s'il est possible d'obtenir le même résultat avec des hypothèses plus faibles, notamment sur les directions et les longueurs des arcs dans Ω_ε. Nous avons supposé que les longueurs des arcs étaient de l'ordre de ε et qu'en tout point $x \in \Omega$, pour une direction admissible v en x, il existait une suite $(y_\varepsilon, e_\varepsilon)_\varepsilon$ qui converge vers (x, v) avec $(y_\varepsilon, e_\varepsilon) \in E^\varepsilon$. Nous pouvons également considérer d'autres modèles discrets. Marcotte [79] s'est intéressé à la question de la programmation à deux niveaux dans les problèmes de congestion. L'équilibre de Wardrop discret dans un cadre markovien (avec les temps qui sont inconnus) a été étudié par Baillon-Cominetti [13]. Dans le modèle continu, nous pouvons également considérer des problèmes de transport optimal dépendant du temps, comme dans les systèmes de jeux à champs moyens (de la théorie de Lasry et Lions [73–75], Mean Field Games en anglais)

2 Régularité des solutions des EDPs

Dans le chapitre 4, nous avons établi un résultat de régularité de Sobolev dans le cas où les directions v_k et les coefficients de volume c_k sont constants. Une question évidente est de savoir si le résultat est encore vrai avec des v_k et/ou des c_k non constants (avec une bonne régularité). Les résultats des simulations numériques réalisées dans ce même chapitre peuvent nous aider à deviner les hypothèses optimales pour avoir le même résultat de régularité pour les solutions de ces EDPs. Il existe de nombreux travaux de Brasco avec des coauteurs sur ces questions de régularité pour ces EDPs dégénérées et elliptiques (qu'elles soient anisotropiques ou non) [28, 31–33]. Nous pouvons également regarder les travaux de DiBenedetto [50], Colombo-Figalli [42], Esposito-Mingione-Trombetti [55] notamment.

3 D'autres applications de l'algorithme ALG2

Comme expliqué dans Benamou-Carlier [21], les méthodes de Lagrangien augmenté conviennent pour traiter des problèmes de MFG. Dans cet article, ils ont utilisé un Hamiltonien quadratique et des coûts de fonctionnement et terminal spécifiques. Dans Buttazzo-Jimenez-Oudet [38], au lieu d'un coût terminal, ils ont considéré une densité terminale. Un modèle plus général de MFG avec diffusion est le système suivant :

$$\begin{cases} \partial_t \phi + \nu \Delta \phi + H(t, x, \nabla \phi) = \alpha(t, x, \varrho), \\ \partial_t \varrho - \nu \Delta \varrho + \operatorname{div}(\varrho \nabla H(t, x, \nabla \phi)) = 0, \\ \varrho|_{t=0} = \varrho_0, \ \phi|_{t=T} = -\gamma(x, \varrho_T) \end{cases}$$

pour un paramètre positif de diffusion ν. Pour des méthodes de résolution numérique basées sur les différences finies appliquées à ce système de MFG avec diffusion, nous pouvons regarder des articles de Achdou avec des coauteurs [1–3]. Une stratégie de Lagrangien augmenté peut également être utilisée mais l'algorithme fait apparaître un opérateur bilaplacien, ce qui engendre des difficultés numériques, comme expliqué dans Achdou-Perez [4].

Bibliographie

[1] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games : numerical methods for the planning problem. *SIAM Journal on Control and Optimization*, 50(1) :77–109, 2012.

[2] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games : convergence of a finite difference method. *SIAM Journal on Numerical Analysis*, 51(5) :2585–2612, 2013.

[3] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games : Numerical methods. *SIAM Journal on Numerical Analysis*, 48(3) :1136–1162, 2010.

[4] Y. Achdou and V. Perez. Iterative strategies for solving linearized discrete mean field games systems. *NHM*, 7(2) :197–217, 2012.

[5] E. Altman, R. El Azouzi, and V. Abramov. Non-cooperative routing in loss networks. *Performance Evaluation*, 49(1) :257–272, 2002.

[6] E. Altman and L. Wynter. Equilibrium, games, and pricing in transportation and telecommunication networks. *Networks and Spatial Economics*, 4(1) :7–21, 2004.

[7] L. Ambrosio. *Lecture notes on optimal transport problems*. Springer, 2003.

[8] L. Ambrosio and G. Crippa. Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields. In *Transport equations and multi-D hyperbolic conservation laws*, pages 3–57. Springer, 2008.

[9] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows : in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[10] L. Ambrosio and A. Pratelli. Existence and stability results in the l^1 theory of optimal transportation. In *Optimal transportation and applications*, pages 123–160. Springer, 2003.

[11] P. Appell. *Mémoire sur les déblais et les remblais des systèmes continus ou discontinus : présenté à l'Académie des Sciences pour le concours du prix Bordin pour 1884*. Académie des Sciences de l'Institut National de France, 1887.

[12] J.-B. Baillon and G. Carlier. From discrete to continuous Wardrop equilibria. *Networks and Heterogenous Media*, 7(2), 2012.

[13] J.-B. Baillon and R. Cominetti. Markovian traffic equilibrium. *Mathematical Programming*, 111(1-2) :33–56, 2008.

[14] D. Bao, S.-S. Chern, and Z. Shen. *An introduction to Riemann-Finsler geometry*, volume 200. Springer Science & Business Media, 2012.

[15] J. W. Barrett and L. Prigozhin. A mixed formulation of the Monge-Kantorovich equations. *M2AN Math. Model. Numer. Anal.*, 41(6) :1041–1060, 2007.

[16] M. Beckmann. A continuous model of transportation. *Econometrica : Journal of the Econometric Society*, pages 643–660, 1952.

[17] M. Beckmann, C. McGuire, and C. Winsten. Studies in the economics of transportation. Technical report, 1956.

[18] M. Belloni and B. Kawohl. The pseudo-p-laplace eigenvalue problem and viscosity solutions as p $\to \infty$. *ESAIM : Control, Optimisation and Calculus of Variations*, 10(01) :28–52, 2004.

[19] J.-D. Benamou. Numerical resolution of an « unbalanced » mass transport problem. *ESAIM : Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 37(5) :851–868, 2003.

[20] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3) :375–393, 2000.

[21] J.-D. Benamou and G. Carlier. Augmented Lagrangian methods for transport optimization, Mean-Field Games and degenerate PDEs. *Journal of Optimization Theory and Applications*, pages 1–26, 2015.

[22] J.-D. Benamou, G. Carlier, and R. Hatchi. A numerical solution to Monge's problem with a Finsler distance as cost. Technical report, 2015.

[23] F. Benmansour, G. Carlier, G. Peyré, and F. Santambrogio. Numerical approximation of continuous traffic congestion equilibria. *Networks and Heterogeneous Media*, 4(3) :605–623, 2009.

[24] F. Benmansour, G. Carlier, G. Peyré, and F. Santambrogio. Derivatives with respect to metrics and applications : subgradient marching algorithm. *Numerische Mathematik*, 116(3) :357–381, 2010.

[25] M. Bernot, J.-M. Morel, and V. Caselles. Traffic plans. In *Publicacions Matemàtiques*, volume 49, pages 417–451, 2005.

[26] G. Bouchitté and G. Buttazzo. Characterization of optimal shapes and masses through Monge-Kantorovich equation. *Journal of the European Mathematical Society*, 3(2) :139–168, 2001.

[27] G. Bouchitté, G. Buttazzo, and P. Seppecher. Mathématiques/mathematics shape optimization solutions via Monge-Kantorovich equation. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 324(10) :1185–1191, 1997.

[28] P. Bousquet, L. Brasco, and V. Julin. Lipschitz regularity for local minimizers of some widely degenerate problems. *arXiv preprint arXiv :1409.2162*, 2014.

[29] A. Braides. *Gamma-convergence for Beginners*, volume 22. Oxford University Press Oxford, 2002.

[30] A. Braides, G. Buttazzo, and I. Fragalà. Riemannian approximation of Finsler metrics. *Journal : Asymptotic Anal*, 31(2) :177–187, 2002.

[31] L. Brasco and G. Carlier. Congested traffic equilibria and degenerate anisotropic PDEs. *Dynamic Games and Applications*, 3(4) :508–522, 2013.

[32] L. Brasco and G. Carlier. On certain anisotropic elliptic equations arising in congested optimal transport : local gradient bounds. *Advances in Calculus of Variations*, 7(3) :379–407, 2014.

[33] L. Brasco, G. Carlier, and F. Santambrogio. Congested traffic dynamics, weak flows and very degenerate elliptic equations. *Journal de mathématiques pures et appliquées*, 93(6) :652–671, 2010.

[34] L. Brasco and M. Petrache. A continuous model of transportation revisited. *Journal of Mathematical Sciences*, 196(2) :119–137, 2014.

[35] Y. Brenier. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *Journal of the American Mathematical Society*, 2(2) :225–255, 1989.

[36] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4) :375–417, 1991.

[37] H. Brezis. *Analyse fonctionelle*, volume 5. Masson, 1983.

[38] G. Buttazzo, C. Jimenez, and E. Oudet. An optimization problem for mass transportation with congested dynamics. *SIAM Journal on Control and Optimization*, 48(3) :1961–1976, 2009.

[39] G. Carlier, C. Jimenez, and F. Santambrogio. Optimal transportation with traffic congestion and Wardrop equilibria. *SIAM Journal on Control and Optimization*, 47(3) :1330–1350, 2008.

[40] G. Carlier and F. Santambrogio. A variational model for urban planning with traffic congestion. *ESAIM : Control, Optimisation and Calculus of Variations*, 11(04) :595–613, 2005.

[41] G. Carlier and F. Santambrogio. A continuous theory of traffic congestion and Wardrop equilibria. *Journal of Mathematical Sciences*, 181(6) :792–804, 2012.

[42] M. Colombo and A. Figalli. Regularity results for very degenerate elliptic equations. *Journal de Mathématiques Pures et Appliquées*, 101(1) :94–117, 2014.

[43] R. Cominetti and J. Correa. Common-lines and passenger assignment in congested transit networks. *Transportation Science*, 35(3) :250–267, 2001.

[44] B. Dacorogna and J. Moser. On a partial differential equation involving the Jacobian determinant. In *Annales de l'Institut Henri Poincaré. Analyse non linéaire*, volume 7, pages 1–26. Elsevier, 1990.

[45] E. De Giorgi. γ-convergenza e G-convergenza. *Ennio De Giorgi*, page 437, 1977.

[46] E. De Giorgi. Convergence problems for functionals and operators. *Ennio De Giorgi*, page 487, 1979.

[47] L De Pascale, L. C. Evans, and A Pratelli. Integral estimates for transport densities. *Bulletin of the London Mathematical Society*, 36(03) :383–395, 2004.

[48] L. De Pascale and A. Pratelli. Regularity properties for monge transport density and for solutions of some shape optimization problem. *Calculus of Variations and Partial Differential Equations*, 14(3) :249–274, 2002.

[49] L. De Pascale and A. Pratelli. Sharp summability for monge transport density via interpolation. *ESAIM : Control, Optimisation and Calculus of Variations*, 10(04) :549–552, 2004.

[50] E. DiBenedetto. $c^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Analysis : Theory, Methods & Applications*, 7(8) :827–850, 1983.

[51] R. Dudley. Probabilities and metrics. *Lecture Notes Series*, 45, 1976.

[52] R. Dudley. *Real analysis and probability*, volume 74. Cambridge University Press, 2002.

[53] J. Eckstein and D. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3) :293–318, 1992.

[54] I. Ekeland and R. Temam. Convex analysis and variational problems. 1976.

[55] L. Esposito, G. Mingione, and C. Trombetti. On the Lipschitz regularity for certain elliptic problems. In *Forum Mathematicum*, volume 18, page 263. WALTER DE GRUYTER & CO., 2006.

[56] L. C. Evans. Partial differential equations and Monge-Kantorovich mass transfer. *Current developments in mathematics*, pages 65–126, 1997.

[57] L. C. Evans and W. Gangbo. *Differential equations methods for the Monge-Kantorovich mass transfer problem*, volume 653. American Mathematical Soc., 1999.

[58] M. Feldman and R. McCann. Uniqueness and transport density in monge's mass transportation problem. *Calculus of Variations and Partial Differential Equations*, 15(1) :81–113, 2002.

[59] M. Fortin and R. Glowinski. Augmented Lagrangian methods, 1983. Applications to the numerical solution of boundary value problems, Translated from the French by B. Hunt and D. C. Spicer.

[60] M. Fukushima. On the dual approach to the traffic assignment problem. *Transportation Research Part B : Methodological*, 18(3) :235–245, 1984.

[61] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1) :17–40, 1976.

[62] W. Gangbo. The monge mass transfer problem and its applications. *Contemporary Mathematics*, 226 :79–104, 1999.

[63] D. Gilbarg and N. Trudinger. *Elliptic partial differential equations of second order*. springer, 2015.

[64] R. Glowinski. An augmented Lagrangian approach to the numerical solution of the Dirichlet problem for the elliptic Monge-Ampere equation in two dimensions. *Electronic Transactions on Numerical Analysis*, 22 :71–96, 2006.

[65] R. Glowinski and A. Marocco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires. *RAIRO 9*, (2) :41–76, 1975.

[66] R. Hatchi. Wardrop equilibria : long-term variant, degenerate anisotropic PDEs and numerical approximations. *arXiv preprint arXiv :1507.01400*, 2015.

[67] R. Hatchi. Wardrop equilibria : rigorous derivation of continuous limits from general networks models. *arXiv preprint arXiv :1506.08960*, 2015.

[68] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4) :251–265, 2012.

[69] R. Huilgol and Z. You. Application of the augmented Lagrangian method to steady pipe flows of Bingham, Casson and Herschel–Bulkley fluids. *Journal of non-newtonian fluid mechanics*, 128(2) :126–143, 2005.

[70] L. V. Kantorovich. On the translocation of masses. In *Dokl. Akad. Nauk SSSR*, volume 37, pages 199–201, 1942.

[71] L. V. Kantorovich. On a problem of Monge. *Journal of Mathematical Sciences (New York)*, 133(1) :15–16, 2004.

[72] M. Krasnoselskii. Topological methods in the theory of nonlinear integral equations. 1964. *Gostekhteoretizdat, Moscow (Eng).*

[73] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9) :619–625, 2006.

[74] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. ii–horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10) :679–684, 2006.

[75] J.-M. Lasry and P.-L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1) :229–260, 2007.

[76] P. Lindqvist. *Notes on the p-Laplace equation.* Univ., 2006.

[77] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*, volume 31. Dunod Paris, 1969.

[78] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6) :964–979, 1979.

[79] P. Marcotte. Network design problem with congestion effects : A case of bilevel programming. *Mathematical Programming*, 34(2) :142–162, 1986.

[80] G. Dal Maso. Introduction. In *An Introduction to* Γ*-Convergence*, pages 1–7. Springer, 1993.

[81] G. Monge. *Mémoire sur la théorie des déblais et des remblais.* De l'Imprimerie Royale, 1781.

[82] J. Moser. On the volume elements on a manifold. *Transactions of the American Mathematical Society*, pages 286–294, 1965.

[83] J. Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[84] N. Papadakis, G. Peyré, and E. Oudet. Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences*, 7(1) :212–238, 2014.

[85] P. Pedregal. *Parametrized measures and variational principles*, volume 28. Birkhäuser, 2012.

[86] R. T. Rockafellar. *Convex analysis.* Princeton university press, 2015.

[87] T. Roughgarden and É. Tardos. How bad is selfish routing ? *Journal of the ACM (JACM)*, 49(2) :236–259, 2002.

[88] T. Roughgarden and É. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2) :389–403, 2004.

[89] F. Santambrogio. Absolute continuity and summability of transport densities : simpler proofs and new estimates. *Calculus of Variations and Partial Differential Equations*, 36(3) :343–354, 2009.

[90] F. Santambrogio. A dacorogna-moser approach to flow decomposition and minimal flow problems. *arXiv preprint arXiv :1310.2738*, 2013.

[91] F. Santambrogio. Optimal transport for applied mathematicians. *by Birkäuser. A non-final version is available at http ://www. math. u-psud. fr/ santambr/OTAMlastversion. pdf*, 2015.

[92] S. K. Smirnov. Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. *Algebra i Analiz*, 5(4) :206–238, 1993.

[93] M. J. Smith. The existence, uniqueness and stability of traffic equilibria. *Transportation Research Part B : Methodological*, 13(4) :295–304, 1979.

[94] C. Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

[95] C. Villani. *Optimal transport : old and new*, volume 338. Springer Science & Business Media, 2008.

[96] M. I. Vishik. Sur la résolution des problèmes aux limites pour des équations paraboliques quasi-linéaires d'ordre quelconque. *Mat. Sbornik*, 59 :289–325, 1962.

[97] M. I. Vishik. Quasi-linear strongly elliptic systems of differential equations of divergence form. *Trudy Moskovskogo matematicheskogo obshchestva*, 12 :125–184, 1963.

[98] J. G. Wardrop. Road paper. some theoretical aspects of road traffic research. In *ICE Proceedings : Engineering Divisions*, volume 1, pages 325–362. Thomas Telford, 1952.

[99] L. Young. *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*. 1937.

[100] L. Young. Generalized surfaces in the calculus of variations. *Annals of mathematics*, pages 84–103, 1942.

[101] L. Young. Generalized surfaces in the calculus of variations II. *Annals of mathematics*, pages 530–544, 1942.